# REPORT DOCUMENTATION PAGE

OMB No. 074-0188



| 1. AGENCY USE ONLY (Leave blank) | 2. REPORT DATE | 3. REPORT TYPE AND DATES COVERED |
|---|---|---|
| | October 1999 | Annual Summary (01 Oct 98 - 30 Sep 99) |

**4. TITLE AND SUBTITLE**
Analysis BAP-1 as a Ubiquitin Hydrolase in the BRCA1 Pathway

**5. FUNDING NUMBERS**
DAMD17-98-1-8269

**6. AUTHOR(S)**
David C. Schultz, Ph.D.

**7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)**
The Wistar Institute
Philadelphia, Pennsylvania 19104

e-mail:
dshultz@wistar.upenn.edu

**8. PERFORMING ORGANIZATION REPORT NUMBER**

**9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)**
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

**10. SPONSORING / MONITORING AGENCY REPORT NUMBER**

**11. SUPPLEMENTARY NOTES**
This report contains colored photos

**12a. DISTRIBUTION / AVAILABILITY STATEMENT**
Approved for public release;
distribution unlimited

**12b. DISTRIBUTION CODE**

**13. ABSTRACT** *(Maximum 200 Words)*

The recurrent theme that has emerged from the structural study of proteins involved in signal transduction and transcriptional regulation is that multiple independently folded globular domains in these proteins often cooperate in macromolecular recognition. These domains are often recognizable by conserved signature amino acid sequence motifs and their spatial organization within a novel protein is often the first clue to its biochemical function. In this aspect, we have identified a novel, nuclear-localized ubiquitin carboxy-terminal hydrolase, BAP-1, that specifically interacts with the RING finger of BRCA1 but not naturally occurring mutants. Biochemical characterization of a second RING finger encoded by a novel transcriptional corepressor protein, KAP-1, has demonstrated that the RING finger in the context of a tripartite RBCC motif (RING-B-box-coiled-coil) is essential for the formation of a structural integrated unit involved in both oligomerization and ligand recognition. Additional studies of the KAP-1 corepressor have shown that KAP-1 directly interacts with the HP1 family of non-histone chromosomal proteins. Interestingly, reduced levels of HP1 have been correlated with aggressive, metastatic behavior of breast carcinoma cell lines. We are currently investigating the biological consequences of restoring wild-type levels of HP1 in these breast cancer cell lines.

**14. SUBJECT TERMS**
Breast Cancer, Transcriptional repression, Heterochromatin, KAP-1, HP-1, KRAB domain

**15. NUMBER OF PAGES**
77

**16. PRICE CODE**

| 17. SECURITY CLASSIFICATION OF REPORT | 18. SECURITY CLASSIFICATION OF THIS PAGE | 19. SECURITY CLASSIFICATION OF ABSTRACT | 20. LIMITATION OF ABSTRACT |
|---|---|---|---|
| Unclassified | Unclassified | Unclassified | Unlimited |

AD_____

Award Number: DAMD17-98-1-8269

TITLE: Analysis BAP-1 as a Ubiquitin Hydrolase in the BRCA1 Pathway

PRINCIPAL INVESTIGATOR: David C. Schultz, Ph.D.

CONTRACTING ORGANIZATION: The Wistar Institute
Philadelphia, Pennsylvania 19104

REPORT DATE: October 1999

TYPE OF REPORT: Annual Summary

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

# FOREWORD

Opinions, interpretations, conclusions and recommendations are those of the author and are not necessarily endorsed by the U.S. Army.

_____ Where copyrighted material is quoted, permission has been obtained to use such material.

_____ Where material from documents designated for limited distribution is quoted, permission has been obtained to use the material.

_____ Citations of commercial organizations and trade names in this report do not constitute an official Department of Army endorsement or approval of the products or services of these organizations.

_✓_ In conducting research using animals, the investigator(s) adhered to the "Guide for the Care and Use of Laboratory Animals," prepared by the Committee on Care and use of Laboratory Animals of the Institute of Laboratory Resources, national Research Council (NIH Publication No. 86-23, Revised 1985).

_✓_ For the protection of human subjects, the investigator(s) adhered to policies of applicable Federal Law 45 CFR 46.

_✓_ In conducting research utilizing recombinant DNA technology, the investigator(s) adhered to current guidelines promulgated by the National Institutes of Health.

_✓_ In the conduct of research utilizing recombinant DNA, the investigator(s) adhered to the NIH Guidelines for Research Involving Recombinant DNA Molecules.

_____ In the conduct of research involving hazardous organisms, the investigator(s) adhered to the CDC-NIH Guide for Biosafety in Microbiological and Biomedical Laboratories.

David C. Schutt 12/23/99

PI Signature Date

# Table of Contents

**Introduction**

The purpose of this post-doctoral fellowship was to further broaden my training in molecular biology/biochemistry, with a particular focus on understanding the biochemical functions of novel gene products, which may have a role in the etiology of human disease. The presence of highly conserved signature amino acid sequence motifs and their spatial organization within a novel protein is often the first clue to its biochemical function. Furthermore, the identification of interacting proteins has been beneficial in defining the biochemical function of a particular protein and/or the cellular pathway (i.e. Cell signalling, transcription, protein degradation) in which it functions. Ultimately, the detailed analysis of a protein's structure, biochemical function, or cellular pathway may assist in the design and development of better therapeutic strategies for the intervention of disease. We have initially used the RING finger motif as a model domain for studying protein function. In particular we have focused our attention on the function of the RING finger in BRCA1. These initial studies have expanded to understanding the function of the RING finger motif in the context of a tripartite RBCC domain (RING-B-box-coiled-coil). A highly conserved spatial arrangement of signature motifs found in a number of oncoproteins, including PML1 and MID1. These studies will have widespread implications in future studies aimed at determining the ligands for RBCC containing proteins. Current efforts are aimed at understanding the biochemical functions of the chromodomain, chromoshadow domain, PHD finger, and bromodomain. Each of these conserved modules is commonly found in proteins with a potential role in the formation and/or maintenance of chromatin structure and function. Throughout the course of these studies a variety of biochemical techniques have been used to investigate the structure-function relationship of these domains. We have extensively charcterized several independent protein-protein interactions between BRCA1 and BAP-1, KAP-1 and a KRAB domain, and KAP-1 and the HP1 protein family. Overall, these studies have increased our understanding of the possible cellular functions of these proteins, and greatly fostered the development of new hypotheses to test experimentally in the future.

**Completed research**

The recurrent theme that has emerged from the structural study of proteins involved in signal transduction and transcriptional regulation is that multiple independently folded globular domains in these proteins often cooperate in macromolecular recognition. We have used the modularity of highly conserved amino acid signature motifs as an approach to studying a protein's structure and function. The RING finger motif is a very abundant sequence motif in the protein database. This highly conserved module of approximately 50 amino acids is defined by eight cysteines and/or histidines that coordinate two zinc ions in a unique cross braced fashion. Upon cloning of *BRCA1*, the most recognizable signature motif in the amino acid sequence was a RING finger at the N-terminus. Moreover, genetic analysis of breast cancer kindreds identified naturally occuring germline mutations that lead to single amino acid substitutions of highly conserved residues in the RING finger and predispose individuals to disease. BAP-1, BRCA1 Associated Protein 1, is a novel, exclusively nuclear ubiquitin carboxy-terminal hydrolase that was identified by the Rauscher lab in a yeast two-hybrid interaction screen for proteins that would interact with the RING finger motif of BRCA1 (Jensen et al. Oncogene 16:1097-1112, 1998). Utilizing a putative coiled-coil structural motif located near the C-terminus of BAP-1, we performed a second, subsequent two-hybrid screen to identify additional proteins that may interact with BAP-1. Sequence from five positive clones, two independent fusions, identically matched nucleotide sequence which encodes the C-terminus of β-catenin. Quantitative β-gal assays of yeast extracts revealed the interaction was weak, indicating that either a larger portion of β-catenin is necessary for optimal interaction or that the interaction might be very transient. To determine if a direct interaction occurred between BAP-1 and β-catenin, $^{35}$S-met labeled *in vitro* translates of BAP-1 and β-catenin were tested for their ability to bind to GST fusions of β-catenin or the C-terminus of BAP-1 *in vitro*, respectively. β-catenin was observed to show some specific association with the C-terminus of BAP-1. In order to verify an *in vivo* interaction between BAP-1 and β-catenin, co-immunoprecipitation experiments were employed. In order to verify an *in vivo* interaction between BAP-1 and β-catenin, co-immunoprecipitation of $^{35}$S-met labeled protein from transiently transfection of CMV driven expression plasmids containing the entire open reading frames of BAP-1 and β-catenin into COS1 cells was employed. All attempts to co-immunoprecipitate BAP1 and β-catenin failed under all experimental conditions (i.e. whole cell lysates vs. nuclear extracts, high salt vs. low salt, different detergents) and washing stringencies. Likewise, immunoprecipitation of endogenous BAP-1 followed by Western analysis with anti-sera against β-catenin failed to detect *in vivo* interaction when performed under similar conditions. Identical

results were obtained when β-catenin immunoprecipitates were Westerned for BAP-1. These data may suggest that this putative interaction is either transient in nature or not completely stable to the experimental conditions we have used.

In addition to being a critical member of cadherin complexes, β-catenin has been shown to be a vital component of a TCF/LEF transcription complex. The literature has well established that cellular levels of β-catenin are regulated via the ubiquitin proteasome pathway. Since BAP-1 has observable ubiquitin hydrolase activity, the effect of BAP-1 on β-catenin regulated transcription (CRT) was investigated through transient transcription assays, utilizing a luciferase construct containing DNA binding sites for TCF4/β-catenin complexes (a kind gift from B. Vogelstein). The effect of BAP-1 on β-gal normalized CRT was less than 2-fold. Moreover, transfection of a catalytically inactive or a C-terminal truncation of BAP-1, which overlaps with the putative β-catenin:BAP-1 interaction domain, demonstrated no specificity for the changes observed in transcription. As a control, transfection of APC was observed to have a 2-5 fold effect on CRT, which was consistent with previously published reports.

The potential effect of BAP-1 on the transcriptional processes of BRCA1 was also analyzed. BRCA1 has been reported to demonstrate p53 independent transcription of WAF1/p21. Since BAP-1 was identified based on its interaction with the RING finger of BRCA1, the effect of BAP-1 on BRCA1 stimulated transcription from a genomic *WAF1* luciferase construct (a kind gift from W. El-Diery) was investigated. However, no induction of luciferase was observed upon transfection of BRCA1 into SW480 cells, a result consistently observed in every cell line transfected (i.e. MCF-7, 293, NIH 3T3). This observation is in complete disagreement with the original published report. Furthermore, it was observed that co-expression of BRCA1 with p53 yielded a 2-fold stimulation of p53 regulated transcription from the WAF1-luciferase construct. This observation is in agreement with two recently published reports. However, expression of BAP-1 was observed to have little to no specific effect on the BRCA1/p53 directed transcription. Moreover, a naturally occurring mutation in the RING finger of BRCA1 (C61G) did not effect the coactivation properties of BRCA1 in this assay, implying that the RING finger or any protein that associates with this motif is not critical to the transcriptional properties of BRCA1 in these particular assays. Independent studies focused at studying the effects of BAP-1 on BRCA1 mediated transcription-coupled repair have been initiated as an alternative system for studying the biochemical functions of BAP-1. Overall, the experimental data described do not provide a convincing argument for the potential interaction between BAP-1 and β-catenin. The presence of

an *in vivo* interaction between BAP-1 and β-catenin could be very transient. This hypothesis is plausible since BAP-1 appears to be an enzyme and may possess quick on/off rates with its substrates. However, if such a protein complex exists, it is likely that BAP-1 does not effect the transcriptional properties of β-catenin. Similarly, it appears that BAP-1 may not effect the transcriptional activation properties of BRCA1, as tested.

It was decided not to continue the analysis of a potential BAP-1/β-catenin interaction, based on the limited preliminary data. Since the goal of this post-doctoral fellowship is to obtain skills that can be used to study the functional relationship between proteins and do to very little preliminary data on the biological function of BAP-1, alternative studies have been pursued. On this line, the Rauscher lab has independently identified a novel corepressor protein KAP-1, KRAB Associated Protein 1 (Friedman et al. Gene Dev. 10:2067-2078, 1997), which also contains a RING finger motif at its amino terminus. KAP-1 functions as a universal corepressor for the superfamily of KRAB-domain zinc finger proteins. The KRAB domain, a 75 amino acid module encoded by approximately one-third of all Kruppel-type C2H2zinc finger proteins, is a modular DNA binding-dependent repression domain, and transcriptional repression directed by a KRAB domain was absolutely dependent upon KAP-1. We have identified several structural elements on KAP-1, of which the biochemical function of each has been thoroughly investigated. In this regard, the interaction between KAP-1 and the KRAB domain has been mapped to overlap with a unique tripartite motif composed of a RING finger, B-box, and coiled-coil region, collectively referred to as the RBCC domain(Peng et al ,in press). This interaction is has been recapitulated from purified components *in vitro*. These experiments indicate that the KRAB:KAP-1 interaction is direct, specific and that the RBCC region is both necessary and sufficient for this interaction. Furthermore, each sub-domain of this tripartite motif is essential for the KRAB:KAP-1 interaction, and each of these structural elements appears to be vital for proper oligomerization of the KAP-1 protein. Biochemical/biophysical experiments indicate that KAP-1 favors an irreversible homo-trimer or homo-hexamer species in solution (Peng, Schultz, and Rauscher, unpublished data). It is not clear from our present studies whether this oligomeric state of KAP-1 is essential for repression of transcription, however, these questions are the focus of current efforts.

We have also conducted studies to explore the repression function of KAP-1. These studies have identified at least three independent repression domains. We have used a variety of biochemical experiments to determine the mechanism of these repression domains and found that one domain directly interacts with members of the HP1 family of heterochromatin proteins, a family non-histone chromosomal protein (Ryan et al, MCB 19: 4366, 1999). This interaction has been recapitulated *in vitro* with purified components, and mapping studies have defined the

minimal interface for interaction in each protein. Specific mutations have been made that disrupt this interaction and demonstrate that this domain is necessary for optimal repression of transcription by KAP-1 *in vivo*. Immunolocalization studies with both monoclonal and polyclonal antibodies specific to KAP-1 indicate that KAP-1 localizes to nuclear domains that are consistent with A-T rich islands of condensed chromatin in interphase nuclei of NIH/3T3 cells. Moreover, KAP-1 staining patterns demonstrate a dynamic staining pattern throughout the nucleus, suggesting that the function of KAP-1 may be regulated in a cell cycle fashion (Schultz, unpublished data). Monoclonal antibodies specific to the three mammalian isoforms of HP1 proteins have been generated to characterize the nuclear localization unique to each protein. Like KAP-1 staining, HP1 alpha staining is concentrated in A-T rich islands of condensed chromatin and appears to be exquisitely coregulated with KAP-1 staining in a cell cycle dependent manner. Staining with antibodies raised against the murine homolog of HP1 beta (M31) displayed a unique staining of condensed foci of chromatin that appear to be juxtaposed to the A-T rich islands of DNA recognized by HP1 alpha antibodies (Schultz, unpublished data). Immunostaining with antibodies specific for HP1 gamma was exclusively nuclear, excluded from the nucleolus, and consistent with euchromatic DNA regions (Schultz, unpublished data). Thus, not all chromodomain containing proteins localize to regions of condensed chromatin in interphase nuclei. Dual staining experiments indicate a high degree of colocalization between KAP-1 and both euchromatic and heterochromatic HP1 proteins in interphase nuclei. Together, these data suggest that KAP-1 may mediate transcription repression via the nucleation and/or maintenance of higher order chromatin structure.

A second autonomous repression domain has been mapped to the PHD finger and bromodomain at the C-terminus of KAP-1. This repression activity depends upon both of these structural elements, as we have found that specific amino acid substitutions in either element impair this repression activity. Furthermore, this repression activity is specific to KAP-1 sequences, as determined by domain swap experiments. These mutations have been used as a tool to identify the potential mechanism(s) of this independent repression domain. A yeast two-hybrid screen revealed several potential downstream effectors including a specific interaction with C-terminal amino acids of the dermatomyositis autoantigen, Mi-2 alpha. This interaction in yeast directly correlated with mutations that disrupt repression activity and was dependent upon the bromodomain. Overexpression of these C-terminal amino acids of Mi-2 alpha can dominantly inhibit KRAB-mediated repression *in vivo*. Since Mi-2 has been biochemically purified as an integral component of a core histone deacetylase complex, these studies suggest that the KRAB:KAP-1 repression complex assists in targeting the biochemical activities of NuRD to specific promoters *in vivo*. Consistent with this hypothesis, addition of trichostatin A to transient transfection assays revealed that the transcriptional repression exerted by the PHD finger and bromodomain of KAP-1 is

partially reversible, indicating the potential involvement of histone deacetylases in KAP-1 mediated repression (Schultz et al., in preparation)

**Future Directions:**

Although the focus of the original proposal has been shifted from characterizing BAP-1, a novel ubiquitin carboxy-terminal hydrolase, in cellular BRCA1 pathways to defining the biochemical properties of the KRAB:KAP-1 repression complex, the experimental approaches are nearly identical. Since the relative abundance of KRAB-ZFP's may potentially make KRAB:KAP-1 mediated repression one of the single most common mechanisms to repress transcription by sequence-specific factors in vertebrates, and the recent observations that that KRAB-ZFP's and members of the KAP-1/TIF1 family of transcriptional regulators may be targets of oncogenic mutations, we will continue to focus our efforts on studying the biochemical properties and biological relevance of the KRAB:KAP-1 repression system.

In this regard, there are several interesting aspects of this current work. The first involves the recent description of RING finger proteins functioning as E3's, ubiqitin ligases, in the ubiquination of proteins *in vitro*. We have initiated a collaboration to test whether the RING finger of KAP-1 or any other members of the TIF1 family of transcriptional regulators can mediate this process. Such studies may provide information as to how this unique family of proteins functions to regulate transcription, and moreover, may provide insights into how these proteins themselves are regulated. In addition, the PHD finger motif, defined by a consensus of C4HC3, significantly resembles the signature sequence of the RING finger. We are currently in the process of determining whether any functional cross-talk between RING fingers and PHD fingers exist. Furthermore, to study the function of the PHD finger in KAP-1 mediated repression we have iniated experiments to both solve the solution structure and crystal structure of the KAP-1 PHD finger. These results will have widespread implications, as more than 40 independent proteins have been described to possess a PHD finger, and mutations in the PHD finger of the hATRX protein predispose individuals to developmental defects.

Finally, it was discovered that levels of HP1 alpha message were significantly reduced in aggressive, metastatic breast carcinoma cell lines. These preliminary observations imply that KRAB-zfp's, KAP-1, or downstream effectors of KRAB:KAP-1 repression may be targets for oncogenesis. We are currently confirming the potential role that reduced expression of HP1 alpha plays in the aggressive, metastatic behavior of these cell lines. We have constructed inducible expression vectors containing the wild-type HP1 alpha cDNA sequence, as well as sequences possessing mutations that produce proteins with well-characterized biochemical effects. We are planning to establish stable cell lines which demonstrate inducible expression of HP1 alpha and its derivatives in order to test the effect that dosage compensation of HP1 alpha may have on reverting

the aggressive behavior of these cell lines. Overall, these new studies are aimed at investigating the role of long term repression mechanisms and their potential role metastatic tumors.

## Appendices

### I. Key Accomplishments

- Identified an *in vivo* interaction between BAP-1 and BRCA1 overexpressed in COS1 cells.
- Identified a direct interaction between the novel corepressor KAP-1 and mammalian members of the HP1 family of non-histone chromosomal proteins.
- KAP-1 and both euchromatic and heterochromatic HP1 proteins can physical occupy the same space in interphase nuclei of NIH/3T3 cells as determined by immunofluoresence staining.
- Generated and characterized both polyclonal and monoclonal antibodies against KAP-1 and each of the mammalian HP1 proteins (alpha, beta, gamma).
- Determined that the tripartite RING-B-box-coiled-coil region of KAP-1 functions as an integrated structural unit in recognition of the KRAB domain and oligomerization of KAP-1.
- Identified a novel DNA-binding dependent repression activity conferred by the PHD finger and bromodomain of KAP-1.
- Used the yeast two-hybrid interaction trap assay to verify the specificity of potential downstream effectors of KAP-1 PHD/bromodomain mediated transcriptional repression, including a potential role for histone deacetylases.

### II. Publications:

Jensen D, Proctor M, Marquis S T, Pery-Gardner H, Ha S I, Ishov A, Tommerup N, Vissing H, Sekido Y, Minna J, Chodosh L A, Borodovsky A, **Schultz D C,** Wilkinson K G, Maul G G, Barlev N, Berger S L, Prendergast G, Rauscher, III F J. BAP1: A Novel Ubiquitin Hydrolase Which Binds to the BRCA1 RING Finger and Enhances BRCA1-Mediated Cell Growth Suppression. Oncogene 16:1097-112 (1998).

Ryan RF, **Schultz DC***, Ayyanathan K, Friedman JR, Fredericks WJ, Singh PB, Rauscher III FJ. KAP-1 Corepressor Protein Interacts and Colocalizes with Heterochromatic and Euchromatic HP1 Proteins: A Potential Role for Kruppel-Associeted Box-Zinc Finger Proteins in Heterchromatic Gene Silencing. Mol. Cell Biol. 19: 4366-4378 (1999).

Li X, Peng H, **Schultz DC**, Lopez-Guisa JM, Rauscher III FJ, and Marmorstein R. Structure-Function Studies of the BTB/POZ Transcriptional Repression Domain from the Promyelocytic Leukemia Zinc Finger Oncoprotein. Cancer Res. 59: 5275-5282 (1999).

Peng H, Begg GE, **Schultz DC**, Friedman JR, Jensen DE, Speicher DW, Rauscher III FJ. Reconstitution of the KRAB-KAP-1 Repressor Complex: A Model System for Defining the

Molecular Anatomy of RING-B-box-Coiled-Coil Domain-Mediated Protein-Protein Interactions. J Mol. Bio. (In press, 2000).

* Indicates equal contribution as first author.

**Abstracts:**

***Schultz DC,** Ryan RF, Freidman JR, Rauscher FJ, III Structure-function analysis of the KRAB:KAP-1 repressor/corepressor system suggests a histone deacetylase independent mechanism of transcriptional repression. Proc. Amer. Assoc Cancer Res. **40**:514 (1999)

Lechner MS, **Schultz DC**, Ayyanathan K, Singh P, Rauscher FJ, III Characterization of the KAP-1 corepressor interaction with HP1 proteins. Proc. Amer. Assoc Cancer Res. **40:**129 (1999).

***Schultz DC,** Lechner MS, Peng H, Ayyanathan K, Patel G, Rauscher FJ, III. A potential role for the KRAB-ZFP:KAP-1 transcriptional repression complex in heterochromatin mediated gene silencing. Eighteenth Summer Symposium in Molecular Biology: Chromatin Structure and DNA Function. The Pennsylvannia State University, July 23, 1999.

***Schultz DC,** Lechner MS, Peng H, Ayyanathan K, Patel G, Rauscher FJ, III. A potential role for the KRAB-ZFP:KAP-1 transcriptional repression complex in heterochromatin mediated gene silencing. Mechanisms of Eukaryotic Transcription: Cold Spring Harbor Laboratories. September 4, 1999.

* Indicates platform presentation given by David C. Schultz

**Manuscripts:**

**Schultz DC**, Ryan RF, Friedman JR, Rauscher III FJ. The PHD finger and Bromodomain of the KAP-1 Corepressor Defines an independent Repression Domain that Interacts with Mi-2 alpha to repress transcription. (In preparation)

Oncogene (1998) 16, 1097–1112


# BAP1: a novel ubiquitin hydrolase which binds to the BRCA1 RING finger and enhances BRCA1-mediated cell growth suppression

David E Jensen[1], Monja Proctor[2], Sandra T Marquis[6], Heather Perry Gardner[6], Seung I Ha[6], Lewis A Chodosh[6], Alexander M Ishov[1], Niels Tommerup[3], Henrik Vissing[4], Yoshitaka Sekido[2], John Minna[2], Anna Borodovsky[5], David C Schultz[1], Keith D Wilkinson[5], Gerd G Maul[1], Nickolai Barlev[1], Shelley L Berger[1], George C Prendergast[1] and Frank J Rauscher III[1]

*[1]The Wistar Institute, 3601 Spruce Street, Philadelphia, Pennsylvania 19104, USA; [2]Hamon Center for Therapeutic Oncology Research, University of Texas Southwestern Medical Center, Dallas, Texas 75235, USA; [3]John F. Kennedy Institute, 2600 Glostrup, Denmark and The Department of Medical Genetics, The Panum Institute, University of Copenhagen, DK2200 Copenhagen, Denmark; [4]Department of Molecular Genetics, Novo Nordisk, DK-2880 Bagsvaerd, Denmark; [5]Department of Biochemistry, Emory University, Atlanta, Georgia 30322, USA; [6]Department of Molecular and Cellular Engineering, University of Pennsylvania School of Medicine, Philadelphia, Pennsylvania 19104, USA*

**We have identified a novel protein, BAP1, which binds to the RING finger domain of the Breast/Ovarian Cancer Susceptibility Gene product, BRCA1. BAP1 is a nuclear-localized, ubiquitin carboxy-terminal hydrolase, suggesting that deubiquitinating enzymes may play a role in BRCA1 function. BAP1 binds to the wild-type BRCA1-RING finger, but not to germline mutants of the BRCA1-RING finger found in breast cancer kindreds. *BAP1* and *BRCA1* are temporally and spatially co-expressed during murine breast development and remodeling, and show overlapping patterns of subnuclear distribution. *BAP1* resides on human chromosome 3p21.3; intragenic homozgyous rearrangements and deletions of BAP1 have been found in lung carcinoma cell lines. BAP1 enhances BRCA1-mediated inhibition of breast cancer cell growth and is the first nuclear-localized ubiquitin carboxy-terminal hydrolase to be identified. *BAP1* may be a new tumor suppressor gene which functions in the BRCA1 growth control pathway.**

**Keywords:** ubiquitin hydrolase; BRCA1; chromosome 3p21.3; RING finger

## Introduction

The cloning of the chromosome 17q21 *BRCA1* breast cancer susceptibility gene is a landmark accomplishment in cancer genetics (Miki *et al.*, 1994). Germline mutations in *BRCA1* appear to account for ~50% of familial breast cancers and essentially all families with 17q21-linked inherited susceptibility to ovarian and breast cancer (Szabo and King, 1995). The importance of this gene is underscored by the fact that kindreds segregating constitutional *BRCA1* mutations show a lifetime risk of 40–50% for ovarian cancer and >80% for breast cancer (Easton *et al.*, 1993, 1995). The classification of *BRCA1* as a highly penetrant, autosomal dominant tumor suppressor gene has been genetically confirmed by the finding of frequent LOH of the wild-type allele in breast tumors from mutation carriers (Hall *et al.*, 1990; Miki *et al.*, 1994; Smith *et al.*, 1992). Surprisingly, *BRCA1* mutations in sporadic breast cancer, including those which show 17 g LOH, have yet to be found and *BRCA1* mutations are extremely rare in sporadic ovarian cancer (Futreal *et al.*, 1994; Merajver *et al.*, 1995).

The *BRCA1* locus spans >100 kb comprising 24 exons (Miki *et al.*, 1994). More than 100 constitutional mutations have been identified in *BRCA1* over the entire length of the gene. Some clustering of these mutations has been seen in populations, and genotype-phenotype correlations have been suggested (FitzGerald *et al.*, 1996; Ford *et al.*, 1994; Muto *et al.*, 1996; Roa *et al.*, 1996; Struewing *et al.*, 1995). The majority of germline mutations result in a truncated BRCA1 protein although recurrent missense mutations resulting in amino acid substitutions in kindreds have also been observed (Couch and Weber, 1996). The heterogeneity of BRCA1 mutant proteins produced by this spectrum of genetic mutations suggests that multiple, independent functions and/or protein–protein interaction surfaces are targets for mutational inactivation. However, the biochemical functions of BRCA1 are largely unknown.

The predominant *BRCA1* mRNA of 8.0 kb encodes a 1863 amino acid protein with only a few sequence motifs suggestive of function (Miki *et al.*, 1994). There are two highly conserved regions. The first is the 100 amino acid N-terminus which encodes a RING finger motif, a domain that is predicted to bind zinc and may be a protein–protein interaction motif (Borden *et al.*, 1995; Lovering *et al.*, 1993). The second region is at the C-terminus which contains an acidic region and two copies of a novel motif, designated the BRCT domain. The BRCT domain is present in a variety of putative cell-cycle related proteins, including RAD9 and 53BP1 (Koonin *et al.*, 1996). The most abundant BRCA1 protein is apparently a ~220 kDa phosphoprotein which is predominantly, but apparently not exclusively, nuclear in subcellular distribution (Chen *et al.*, 1995, 1996b; Scully *et al.*, 1996). BRCA1 is localized to discrete nuclear dot structures in a cell-cycle-dependent manner (Scully *et al.*, 1997b). Other isoforms of BRCA1 have been detected including a protein of 97 kDa. This smaller form lacks exon 11, and thus a functional nuclear localization signal, and is presumably the result of an alternative splicing event (Thakur

Correspondence: FJ Rauscher, III
Received 16 January 1998; revised 26 January 1998; accepted 27 January 1998

*et al.*, 1997). The above observations, coupled with the finding of a BRCA1 COOH-terminal domain capable of activating transcription as a Gal4 DNA-binding domain fusion (Chapman and Verma, 1996) and the co-fractionation of BRCA1 with the RNA pol II holoenzyme (Scully *et al.*, 1997a), suggest a role for BRCA1 in transcriptional regulation.

The expression patterns of *BRCA1* further support its role in growth regulation and/or differentiation. The spatial-temporal expression pattern in the embryonic mouse includes the neuroepithelium, and epithelial lineages of the skin, kidney and mammary gland (Marquis *et al.*, 1995). Moreover, *BRCA1* mRNA is sharply increased in alveolar and ductal cells of the breast epithelia during pregnancy (Marquis *et al.*, 1995). Consistent with this, *BRCA1* transcription is under (indirect) hormonal control in both cell culture and organismal systems (Gudas *et al.*, 1995, 1996; Vaughn *et al.*, 1996; Marks *et al.*, 1997). *BRCA1* is also highly expressed in the adult testis during the final stages of meiosis and spermiogenesis (Zabludoff *et al.*, 1996). Together, these observations suggest a broad role for BRCA1 in terminal differentiation events in multiple tissues. Somewhat paradoxically, the murine *brca1−/−* embryos die very early in gestation and exhibit severe cell proliferation defects and profound cell cycle arrest (Hakem *et al.*, 1996; Liu *et al.*, 1996). The association of *BRCA1* expression with both proliferation and differentiation events suggests a possible role for BRCA1 in regulating a genetic program which prepares the cell for terminal differentiation and possibly maintains that phenotype. Results of cell culture and transfection studies have underscored the tumor suppression function of BRCA1, but have revealed little of possible mechanisms. *BRCA1* antisense expression can transform fibroblasts and accelerates growth of breast cancer cell lines (Rao *et al.*, 1996; Thompson *et al.*, 1995). Expression of wild-type *BRCA1* inhibits colony formation and tumor growth *in vivo*, whereas tumor derived mutants of *BRCA1* lack this growth suppression activity (Holt *et al.*, 1996).

Evidence of a role for BRCA1 as a terminal differentiation checkpoint has recently been provided by the finding that BRCA1 and the RAD51 protein (involved in DNA recombination/repair) are co-localized and physically associated in mitotic and meiotic cells (Scully *et al.*, 1997b). The co-localization of BRCA1 and RAD51 on synaptonemal meiotic chromosomes suggests a role for this complex in either the fidelity of DNA replication, cell-cycle progression or genomic integrity. Though intriguing, these results do not suggest a function for BRCA1 which, when lost through mutation of the BRCA1 gene, would give rise to tumors. Strategies based upon identification of proteins which bind to BRCA1 have yielded components of the nuclear import pathway (Chen *et al.*, 1996a) and a novel RING finger/BRCT-domain-containing protein, BARD1 (Jin *et al.*, 1997; Wu *et al.*, 1996). However, none of these associated proteins have suggested a function for BRCA1.

We have chosen to focus upon the highly conserved BRCA1 RING finger domain as a potential protein–protein interface. This motif is defined by a spatially conserved set of cysteine–histidine residues of the form $C_3HC_4$. Structural analysis of the motif shows that two molecules of zinc are chelated by the consensus residues in a unique 'cross-braced' fashion (for reviews, see; Klug and Schwabe, 1995; Saurin *et al.*, 1996). Comparative structure analyses suggest that the RING fingers have a common hydrophobic core structure but that the region encoded by amino acids spanning cysteines 24 and 64 (for BRCA1) forms a highly variable loop structure which may be the determinant of protein–protein interaction specificity. The RING motif occurs in over 80 proteins including the products of proto-oncogenes and putative transcription factors (Saurin *et al.*, 1996). Evidence that the RING finger domain functions as a protein–protein interface has come from the study of the proto-oncogene PML (Borden *et al.*, 1995) and the transcriptional co-repressor KAP-1 (Friedman *et al.*, 1996). Intriguingly, like BRCA1, both PML and KAP-1 are localized to discrete, non-overlapping, nuclear dot structures, and mutations in the RING finger of PML abolish its localization to these dot structures (Borden *et al.*, 1995).

We hypothesize that the BRCA1 RING finger is a binding site for protein(s) which either mediate BRCA1 tumor suppressor function or serve to regulate these functions. Genetic evidence supports this in that single amino-acid substitutions at metal chelating cysteines, C61G and C64G, occur in *BRCA1* kindreds; these mutations segregate with the disease susceptibility phenotype and are predicted to abolish RING finger structure. We have used the yeast two-hybrid system to isolate proteins which bind to the wild-type BRCA1 RING finger but not to the C61G or C64G mutated RING fingers or other closely related RING fingers. We have isolated mouse and human clones of a novel protein, *B*RCA1 *a*ssociated *p*rotein-1 (BAP), which fulfils these criteria. BAP1 is a novel, nuclear-localized enzyme which displays the signature motifs and activities of a ubiquitin carboxy-terminal hydrolase. Full-length BRCA1 binds to BAP1 *in vitro* and enhances the growth suppression properties of BRCA1 in colony formation assays. The human *BAP1* locus was mapped to chromosome 3p21.3, and homozygous deletions of *BAP1* were found in non-small cell lung cancers. Together, these data suggest that *BAP1* is a key player in the *BRCA1* growth suppression pathway, and may itself be a tumor suppressor gene. The identification of BAP1 as a ubiquitin hydrolase implicates the ubiquitin-proteasome pathway in either the regulation, or as a direct effector, of BRCA1 function. BAP1 is the first nuclear-localized ubiquitin carboxy-terminal hydrolase to be identified, and may play a broad role in ubiquitin-dependent regulatory processes within the nucleus, including the emerging roles of ubiquitin conjugation as a post-translational modification which alters protein function and/or subcellular targeting.

## Results

### *A yeast two-hybrid screen for BRCA1 RING finger interacting proteins*

We constructed a synthetic *BRCA1* gene encoding the amino-terminal 100 amino acids of human BRCA1 using long oligonucleotides and PCR-mediated over-

lap-extension gene synthesis techniques (Madden *et al.*, 1991). Codon usage was optimized for expression in *E. coli* and *S. cerevisiae*. The resulting gene was fused to the LexA DNA binding domain (Figure 1b). The negative control/specificity controls included: (1) the Cys61Gly, Cys64Gly, Met18Thr and Arg71Gly muta-

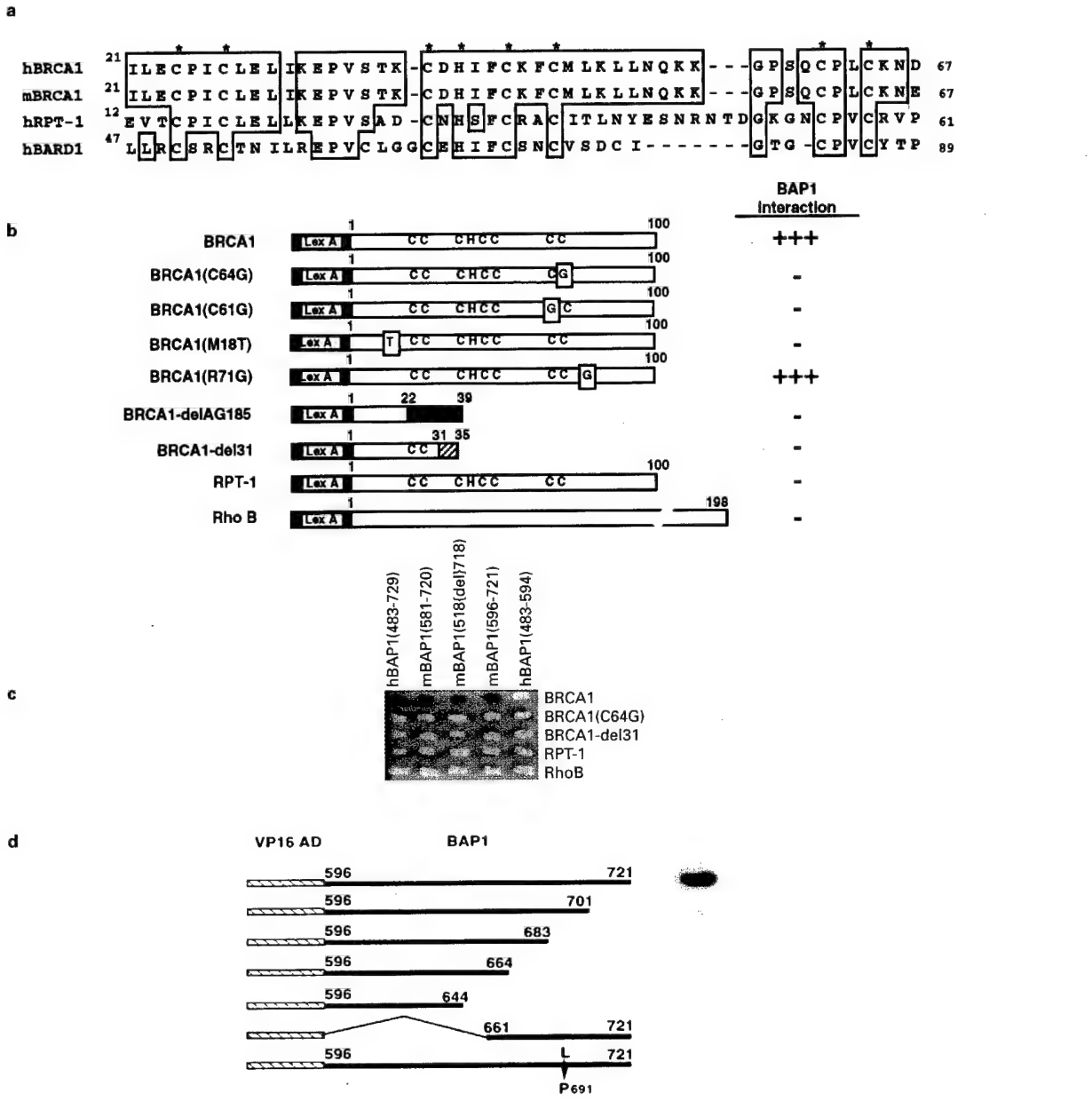
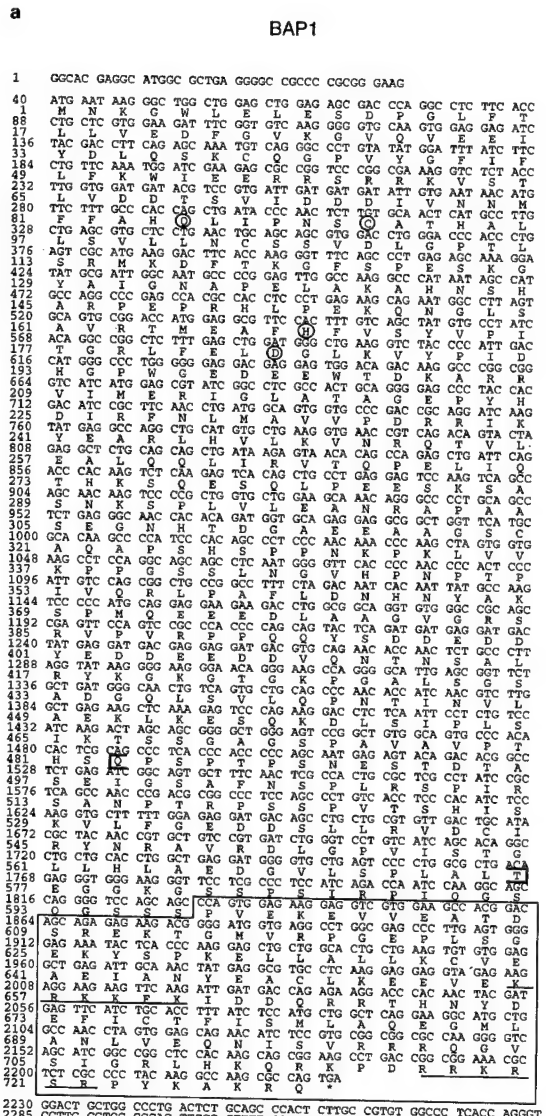
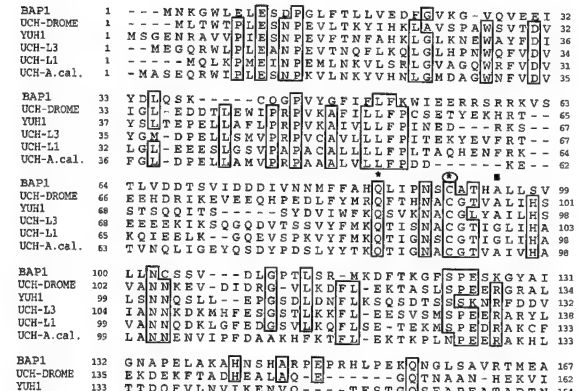



**Figure 1** BRCA1 RING finger domain binds a novel protein. (**a**) Alignment of the RING finger domains of human and mouse BRCA1 (AA's 21–67), RPT-1 (AA's 12–61; the most closely related RING finger) and BARD1 (AA's 47–89). Asterisks (*) identify the Zn-chelating amino acids that form the core of the RING finger. Boxed amino acids show regions of identity between the RING finger domains of human BRCA1 and the other proteins. Alignment performed by CLUSTALW (Thompson *et al.*, 1994). (**b**) The amino-terminal 100 amino acids of human BRCA1 (which includes the RING finger domain) or the indicated amino acids of the various BRCA1–RF mutants and controls were fused to the LexA DNA-binding domain. Expression of all fusions in yeast was confirmed by Western analysis. A summary of the two-hybrid interaction between the Gal4-hBAP1(483–729) fusion clone and the various LexA-RING finger fusions is shown. (**c**) The BRCA1-interacting protein specifically interacts with the BRCA1 RING finger domain. Two hybrid screens of a human B-cell library and a mouse embryo (9.5–10.5 days) library identified a protein that interacted with wild type BRCA1–RF, but not with BRCA1-del31 (a truncated BRCA1), BRCA1(Cys64Gly) (a BRCA1–RF containing a point mutation), RPT-1 (a RING finger closely resembling the BRCA1), or RhoB (a non-related protein). Dark color of yeast indicates transcription from the LacZ reporter gene. Clones obtained from the two libraries are described as partial BAP1 proteins with AAs in parentheses. h, human; m, mouse. (**d**) The two-hybrid interaction between the BRCA1 RING domain and BAP1 requires the BAP1 C-terminal domain. Murine clone mBAP1(596–721) defines a portion of the BRCA1-interaction domain of BAP1. Mutants of this clone were generated by PCR-based deletion or point mutagenesis of mBAP1(596–721) as described in Materials and methods. Each individual mutant was co-transformed with LexA-BRCA1–RF into yeast and tested for interaction via its ability to activate transcription from the LacZ locus. Expression of all fusions in yeast were confirmed by Western analysis

tions that occur in *BRCA1* families; (2) the protein equivalent of the del AG185 mutation which results in a frame-shift at amino acid 22 followed by 17 out-of-frame amino acids and a stop codon; (3) a truncated BRCA1 RING finger at amino acid 31, the result of a PCR error; (4) the RPT-1 RING finger domain, a putative lymphocyte-specific transcription factor, whose RING finger domain is most highly related to



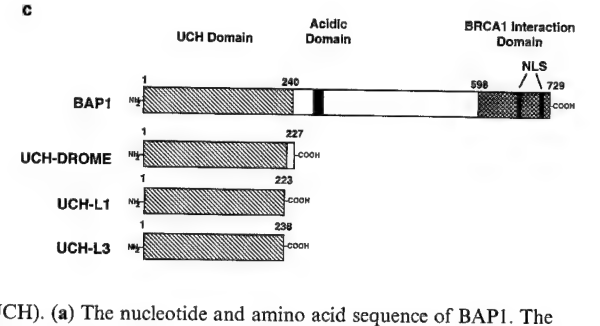




**Figure 2** BAP1 is a novel ubiquitin carboxy-terminal hydrolase (UCH). (**a**) The nucleotide and amino acid sequence of BAP1. The longest open reading frame which contained the amino acids defined by the human 2-hybrid fusion protein is 2188 nucleotides encoding 729 amino acids. The cDNA also contains 39 nucleotides of 5′UTR and 1295 nucleotides of 3′UTR. The enzymatic active site is contained within the first 250 amino acids; the active site residues are circled. The putative nuclear localization signals (NLS) are underlined, the BRCA1-interaction domain is boxed and the protein fragment used to generate BAP1 polyclonal antibodies is bracketed (AAs 483–576). (**b**) Comparison of BAP1 with other UCH's. UCH DROME (genebank #P35122), YUH1 (genebank #P35127), UCHL-1 (genebank #PO9936), UCHL-3 (genebank #P15374), UCH A.cal. (genebank #U90177). The BLAST search algorithm was used to identify proteins closely related to BAP1 (Altschul *et al.*, 1990). The UCH domain of five of these proteins were aligned with BAP1 using the CLUSTALW (ver.1.6) algorithm (Thompson *et al.*, 1994). Residues conserved between species with an identity of 85% or greater are boxed. Active site residues are denoted with asterisks. The catalytic cysteine is marked with a circled asterisk. (**c**) Diagrammatic comparison between BAP1 and some of the other UCH family members

that of BRCA1 (Patarca *et al.*, 1988); and (5) a LexA fusion with RhoB. The wild-type BRCA1 RING finger (BRCA1–RF) did not display intrinsic transcriptional activation function in yeast, and expression of each LexA fusion in yeast was confirmed by Western blot analysis with anti-LexA antibody (data not shown).

Guided by the expression patterns of BRCA1 during mouse development and in human spleen, we chose to screen cDNA libraries constructed from E9.5-10.5d whole mouse embryos and human adult B cells with the LexA-BRCA1–RF. Thirty-one cDNAs which specifically interacted with BRCA–RF were obtained: eight of these (three from the human library and five from the mouse library) encoded the same amino acid sequence. These were designated *BR*CA1 *A*ssociated *P*rotein-1 (BAP1) and pursued further. Each clone shares the same translational reading frame with respect to the transcriptional activation domain to which it is fused. In addition, the fusion junctions were different among the clones, suggesting that the interaction was not due to a fusion-junction artifact. Furthermore, the hBAP1 (483–729) and the BRCA1–RF interacted strongly in a mammalian two-hybrid assay (data not shown). The longest *BAP1* cDNA retrieved in the two-hybrid screen was a ~2.0 kbp clone from the human library and encoded 246 amino acids followed by a 1.3 kb 3′UTR. Each murine clone encoded an overlapping, smaller subset of this human open reading frame with a human-murine AA sequence identity of 100% over the COOH-terminal 100 AAs (data not shown). Both human and mouse clones showed a strong interaction with the wild-type BRCA1–RF and BRCA1 (Arg71Gly) substitution, but failed to interact with the C64G, C61G, del31, delAG, RPT-1, RhoB, or a variety of other LexA fusion constructs (Figure 1b,c and data not shown).

Further definition of this highly conserved interaction domain was performed by mutagenesis of this region of BAP1. Deletion of protein sequence from the carboxyl or amino termini of mBAP1(596–721) almost completely destroyed the BAP1-BRCA1 interaction (Figure 1d), possibly suggesting an extended interface between the proteins. Interestingly, the mBAP1(518del718) clone interacted most poorly with BRCA1–RF (Figure 1c) and lacked a 93 bp sequence (the reading frame was maintained), possibly the result of a naturally occurring splice variant. That BAP1 also failed to bind multiple, independent tumor-derived mutations of the BRCA1–RF provides strong evidence for the relevance of this interaction to the functions of BRCA1.

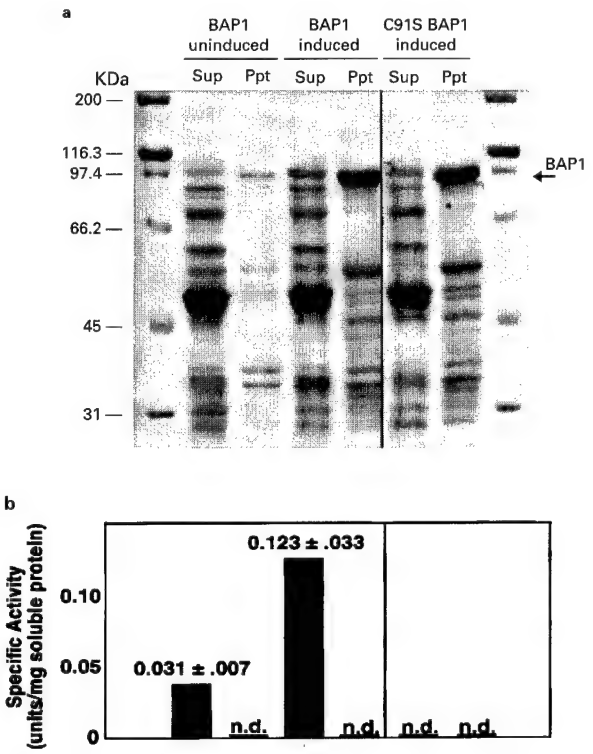



**Figure 3** BAP1 has ubiquitin hydrolase activity. BAP1, or an enzymatically null mutant, BAP1(C91S), were expressed in bacteria by IPTG induction. Bacteria were harvested, lysed and supernatant (Sup) and pellet (Ppt) fractions generated. Each fraction was then measured for UCH activity (bar diagram; n.d., not detected). Induction of protein was verified by SDS–PAGE of each fraction. Arrow indicates BAP1 and BAP1(C91S) protein

*Analysis of the BAP1 cDNA*

A full-length cDNA was constructed using two IMAGE consortium EST clones and RT–PCR (Figure 2a; see Materials and methods). The *BAP1* cDNA comprises 3525 bp; a polyA tract is present along with multiple polyA signals. Conceptual translation yields a long open reading frame of 729 amino acids with a predicted MW of 81 kDa and pI of 6.3. The presumptive initiator methionine is within a favorable context for translation start, however the short 5′UTR of 39 bp encodes amino acids in-frame with the presumptive methionine and does not contain a stop codon. Computer database searches indicated that BAP1 is a novel protein with the amino-terminal 240 amino acids showing significant homology to a class of thiol proteases, designated ubiquitin C-terminal hydrolases (UCH), which are implicated in the proteolytic processing of ubiquitin (Wilkinson *et al.*, 1989). These enzymes play a key role in protein degradation via the ubiquitin-dependent proteasome pathway. Similarities to other mammalian UCHs (UCH-L3 and UCH-L1) have been found (Figure 2b and c). Most importantly, the residues which form the catalytic site (Q85, C91, H169 and D184) are completely conserved, including the FELDG motif (Larsen *et al.*, 1996). In addition, a loop of highly variable sequence, which is disordered in the crystallographic structure of human UCH-L3 (Johnston *et al.*, 1997), is present (residues 140–167). This loop may occlude the active site or provide substrates specificity for the enzyme.

BAP1 has a number of additional motifs; a region of extreme acidity spanning amino acids 396 to 408, as well as multiple potential phosphorylation sites and N-linked glycosylation sites (Figure 2a). The C-terminal one-third is highly charged and is rich in proline, serine and threonine. The extreme C-terminus contains two putative nuclear localization signals, KRKKFK and RRKRSR and is hydrophilic, it is predicted to fold into a helical (possibly coiled-coil) structure (Figure 2a;

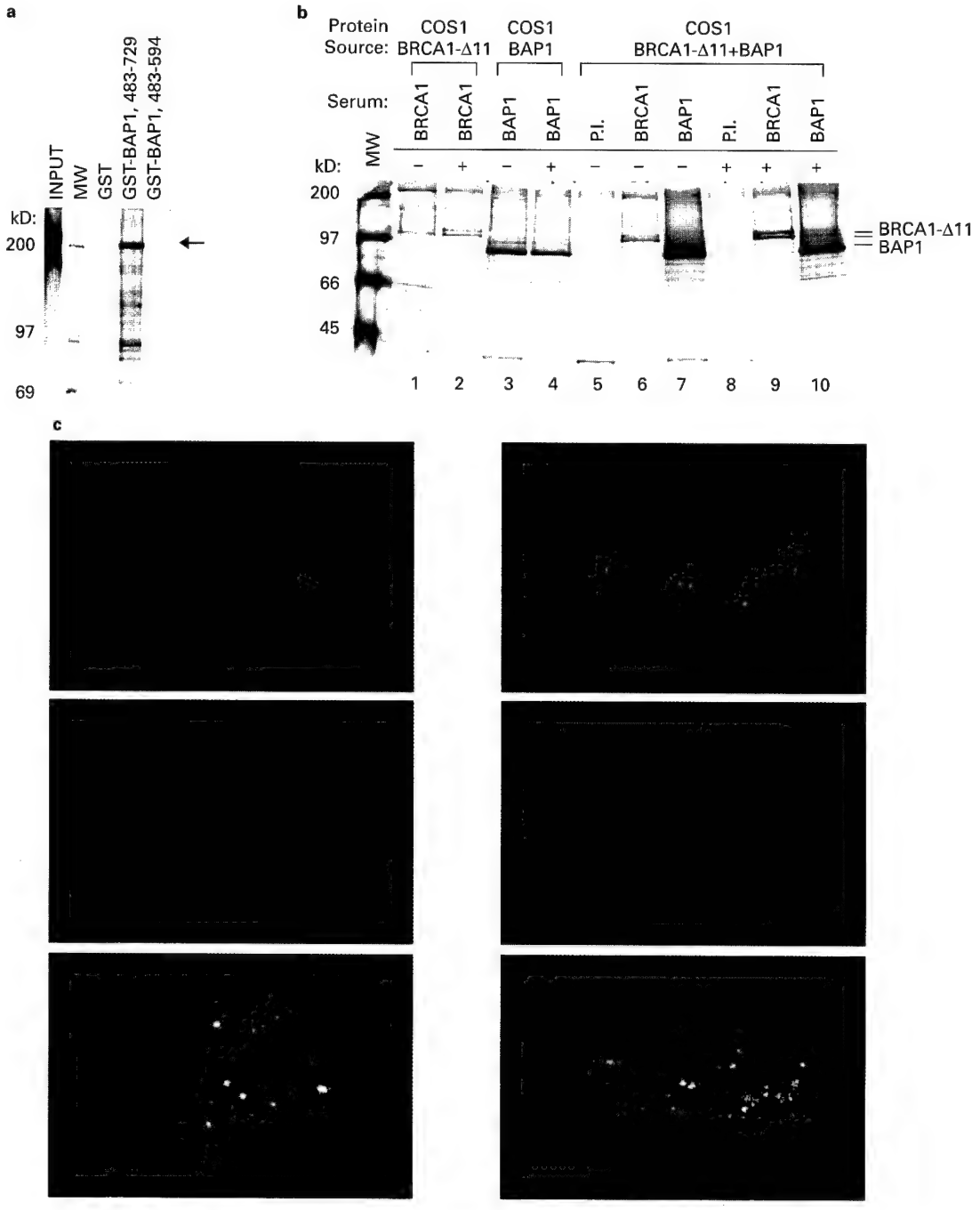



**Figure 4** BAP1 and BRCA1 interact *in vitro* and *in vivo*. (**a**) BRCA1 and GST-hBAP1(483–729) interact *in vitro*. GST, GST-hBAP1(483–729) and GST-hBAP1(483–594) fusion proteins were expressed in *E. coli* and purified as described in Materials and methods. The Glutathione-Sepharose resins, containing an equivalent amount of fusion protein, were incubated in batch with *in vitro*-expressed, $^{35}$S-labeled, BRCA1. After extensive washing, the proteins which remained bound were analysed by SDS–PAGE and fluorography. Lane 1, Input 2% of the labeled BRCA1 used in the associations with the glutathione resins. Lane 2, molecular weight markers. Lane 3, proteins bound to GST alone. Lane 4, proteins bound to GST-hBAP-1(483–729). Lane 5, proteins bound to GST-hBAP1-(483–594), a fusion protein lacking the BRCA1-interaction domain. Arrow indicates the BRCA1 protein. (**b**) Co-immunoprecipitation of BRCA1-Δ11 and BAP1. COS1 cells, transiently transfected with BRCA1-Δ11, BAP1, or BRCA1-Δ11 and BAP1 together, were treated with DMSO (−) or with 25 $\mu$M ALLN (+) for 16 h. Whole cell extracts were prepared after metabolic labeling and then immunoprecipitated with anti-BAP1 antibody (BAP1) or with anti-BRCA1 antibody, C20 (BRCA1) as indicated. Immunoprecipitated proteins were visualized by SDS–PAGE and fluorography. The BAP1 and BRCA1-Δ11 proteins are indicated. (**c**) BRCA1 and BAP1 partially co-localize in the nucleus. Asynchronous Rh30 cells were co-stained with the BRCA1-specific BR64 monoclonal antibody (green, top and bottom panels; Upstate Biotechnology) and a BAP1 specific, affinity-purified polyclonal antibody (red, middle and bottom panels). Top and middle panels are single color readouts of the dual color bottom panel. Yellow is produced where green and red images overlap

S Subbiah, personal communication). Indeed, within this domain the mutation of leucine 691 to a proline, a change predicted to disrupt the helical nature of this region, abolished the BAP1–BRCA1 interaction (Figure 1d). This result is consistent with the hypothesis that BAP1 uses a coiled-coil domain to interact with the RING finger domain of BRCA1. This overall architecture suggests that BAP1 is a new, structurally complex, nuclear-localized member of the UCH enzyme family.

### *BAP1 has UCH activity*

To determine whether BAP1 did indeed have UCH enzymatic activity, BAP1 was expressed in bacteria and this protein was assayed for the ability to hydrolyze the glycine 76 ethyl ester of ubiquitin (Ub-OEt; Mayer and Wilkinson, 1989). IPTG-induced expression of BAP1 in bacteria led to abundant protein, most of which was found in an inactive, insoluble form (Figure 3a, 'BAP1 induced-Ppt'). The BAP1 protein found in the soluble

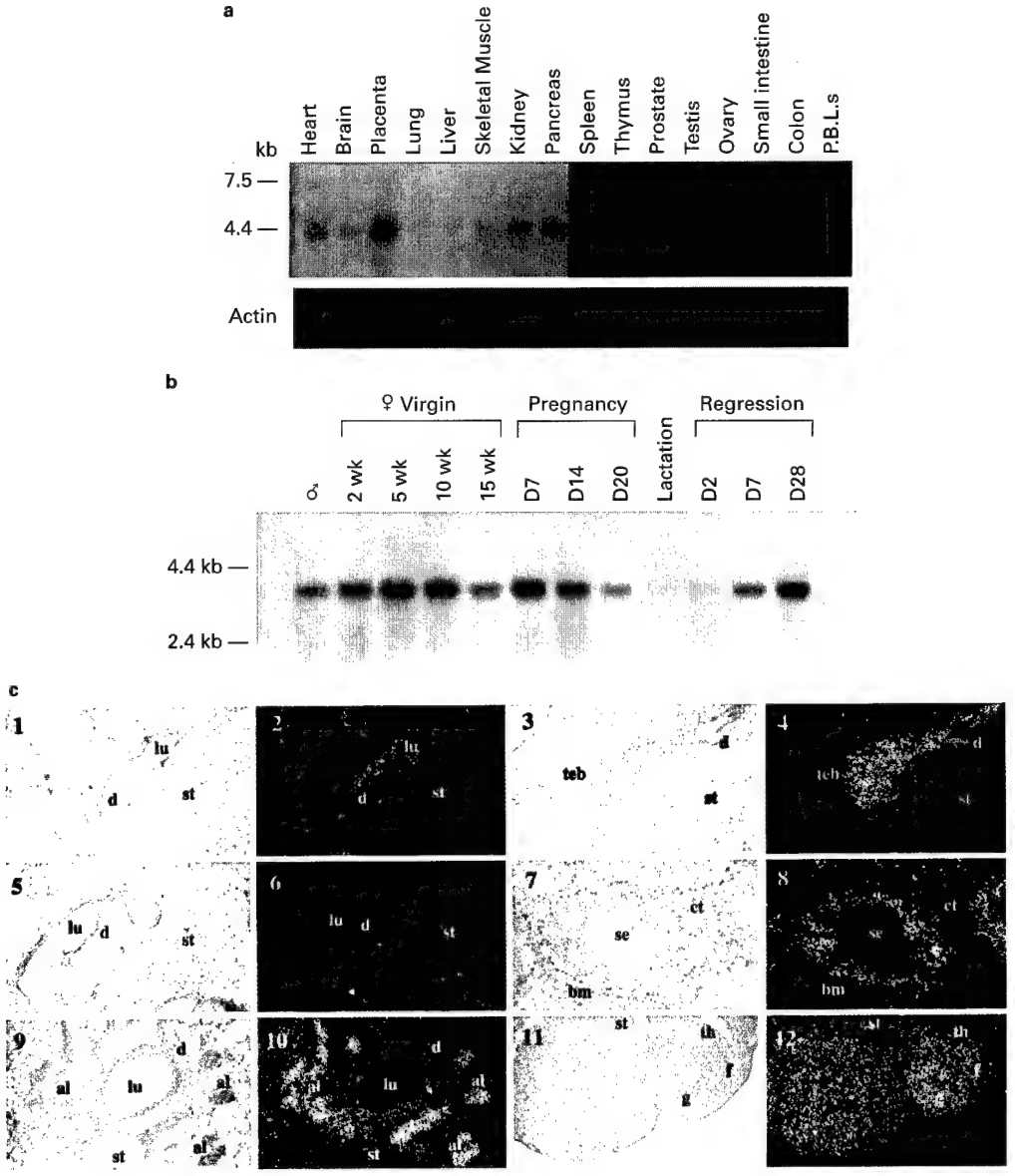



**Figure 5** Tissue, spatial and temporal pattern of BAP1 expression. (**a**) Northern hybridization analysis of human RNA from multiple tissues. Northern blots that contain human RNA from the indicated tissues were probed with $^{32}$P-labeled hBAP-1(483–729) cDNA (nts. 1488–3525). The blots were subsequently probed with a muscle actin cDNA. (**b**) Northern hybridization analysis of *Bap1* expression during mammary gland development. A Northern blot containing poly(A)$^+$ RNA isolated from mouse mammary glands at the indicated developmental stages was probed with $^{32}$P-labeled mBAP1(596–721) cDNA. Note that mouse *Bap1* RNA is slightly smaller than the human *BAP1* RNA. (**c**) *In situ* hybridization analysis of *Bap1* expression. Bright-field (1, 3, 5, 7, 9 and 11) and dark-field (2, 4, 6, 8, 10 and 12) photomicrographs of *in situ* hybridization analyses performed on paraffin sections of mammary glands harvested from adolescent (1–4), mature (5, 6) or pregnant (9, 10) mice. Also shown are sections of testis (7, 8) and ovary (11, 12). To facilitate comparison, dark-field photomicrographs of breast were taken using identical shutter exposure times. al, alveolar bud; d, ductal epithelium; lu, ductal lumen; st, stroma; teb, terminal end bud; d, bm, basement membrane; se, semeniferous tubule; f, developing follicle; g, granulosa cells; th, thecal cells; ct, connective tissue

fraction was able to hydrolyze Ub-OEt, indicating that BAP1 contains UCH-like enzymatic activity (Figure 3b). The active site thiol residue responsible for UCH activity in UCH-L3 has been identified and its mutation leads to abolition of enzyme activity (Larsen *et al*., 1996). Mutation of the corresponding cysteine residue in BAP1, BAP1(C91S), yielded a protein with no detectable UCH activity (Figure 3b) further supporting the conclusion that BAP1 is a thiol protease of the UCH family.

### *BAP1 associates with BRCA1* in vitro *and* in vivo

Association of BRCA1 with BAP1 was tested *in vitro* by binding of full-length BRCA1 to hBAP1(483–729) fused to glutathione S-transferase (GST; Figure 4a). The $^{35}$S-labeled BRCA1, produced by coupled *in vitro* transcription and translation, specifically bound to the GST–hBAP1(483–729) fusion protein, but not to GST alone (Figure 4a) indicating a physical association between the two proteins. As predicted from the yeast two-hybrid results (Figures 1d and 4a), BRCA1 did not bind to GST–hBAP1(483–594), a GST–BAP1 fusion protein lacking the BRCA1 interaction domain.

To determine whether BRCA1 and BAP1 could interact in mammalian cells, a co-immunoprecipitation analysis from co-transfected COS1 cells was performed (Figure 4b). Several attempts to transiently express BRCA1 to any significant level in COS1 cells were without success. Therefore, we performed the analysis with BRCA1-Δ11, a naturally occurring splice variant (Thakur *et al*., 1997) which can be expressed in these cells and which contains the RING finger domain. The proteasome/caspase inhibitor ALLN (N-acetyl-L-Leucinyl-L-Leucinyl-L-norLeucinal) was included in the analysis to determine its influence on the stability of the interaction between BRCA1 and BAP1. BRCA1-Δ11 was detected by immunoprecipitation as a sharp band at ~99 kDa in singly transfected COS1 cells (Figure 4b, lane 1). Incubation of a parallel set of transfected cells with ALLN (20 h) prior to harvest revealed a discrete slower migrating band in the anti-BRCA1 immunoprecipitates (Figure 4b, lane 2). Immunoprecipitates from BAP1-transfected COS1 cells revealed a 91 kDa protein whose mobility was apparently unaffected by treatment with ALLN (Figure 4b, lanes 3 and 4). Co-transfection of COS1 cells with BRCA1-Δ11 and BAP1 revealed that these proteins could be co-immunoprecipitated using either anti-BAP1 or anti-BRCA1 antibodies (Figure 4b, lanes 9 and 10); the ability to detect the BRCA1–BAP1 complex under these conditions was dependent upon incubation of the cells with the proteasome inhibitor. In co-transfected, ALLN-treated cells, both forms of BRCA1-Δ11 are evident in the co-immunoprecipitated complex, and both forms are more abundant than singly transfected cells (compare lanes 2 and 10). These results demonstrate the *in vivo* association of BRCA1 and BAP1, and suggest the presence of a proteasome inhibitor-sensitive modification of BRCA1 which may enhance its interaction with BAP1.

To further document the *in vivo* interaction between BAP1 and full-length BRCA1, we determined whether the endogenous (non-transfected) proteins were co-localized in the cell nucleus (Figure 4c). A mouse monoclonal antibody to BRCA1 (Maul *et al*., manuscript in preparation) and affinity purified rabbit BAP1 antibody were used to stain rhabdomyosarcoma cells (Rh30), a cell line previously determined to express *BAP1* RNA (Jensen and Rauscher, unpublished results). BRCA1 was detected exclusively in punctate domains within the nucleus (excluding nucleoli) in agreement with other reports (Jin *et al*., 1997; Scully *et al*., 1997b). BAP1 was also detected in punctate domains within the nucleus of Rh30 cells, however, the number of domains was significantly greater than for BRCA1. Several of the BRCA1 domains coincided with BAP1 domains (yellow dots in the bottom panels), suggesting an endogenous BRCA1-BAP1 interaction. We also detected BRCA1 reactive domains which were not co-localized with BAP1 domains as has been seen with other BRCA1-interacting proteins (Figure 4c; Jin *et al*., 1997; Scully *et al*., 1997b). These data show that BAP1 and BRCA1 can physically interact *in vitro* and *in vivo*, and have overlapping subnuclear expression patterns.

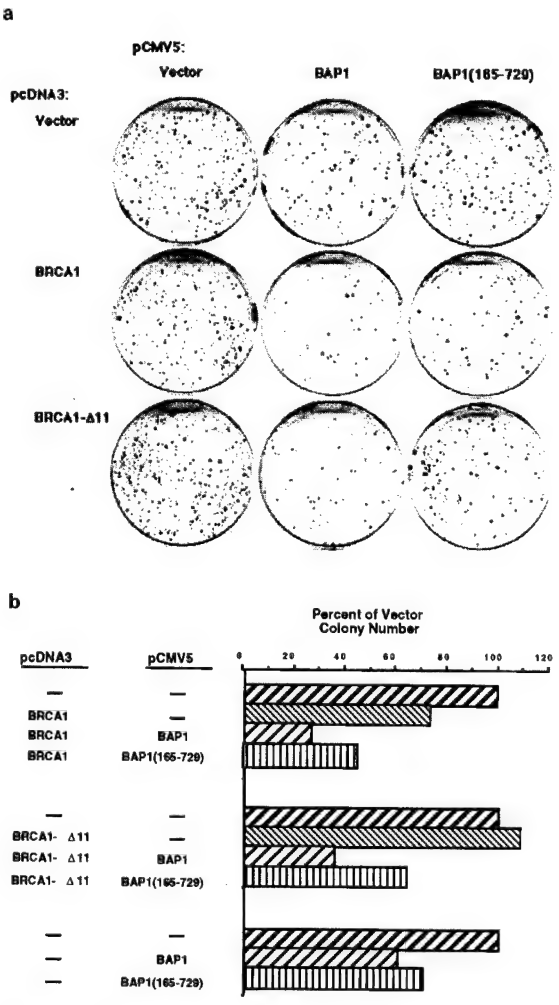



**Figure 6** BAP1 enhances BRCA1-mediated growth suppression. (**a**) MCF7 cells were co-transfected with each of the plasmid constructs shown. Cells were then harvested and $5 \times 10^5$ cells were plated in duplicate into complete medium containing G418. Twenty-one to 28 days later, cells were stained and colonies counted. The experiment was repeated four times with similar results. (**b**) Quantitation of the results from (**a**)

### *BAP1 is expressed in a temporal and spatial pattern during breast development/remodeling*

The physical interaction between BAP1 and BRCA1 suggests that the proteins might be expressed in similar tissues. Northern blot hybridization analysis of *BAP1* expression in a variety of human adult tissues indicated that human *BAP1* was encoded by a single mRNA species of ~4 kb in all tissues except testis, where a second, ~4.8 kb mRNA, was also detected (Figure 5a). High expression was detected in testis, placenta and ovary, with varying levels detected in the remaining tissues. Expression of *BAP1* in normal human breast tissue was detected by RT–PCR of total RNA isolated from normal human mammary epithelial cells (data not shown).

Northern analysis and *in situ* hybridization were performed to determine whether the spatial and temporal pattern of *Bap1* expression in the breast corresponded to that previously described for *Brca1* (Figure 5b and c; Marquis *et al*., 1995). *Bap1* was expressed at slightly higher levels in the mammary glands of 5 week-old adolescent female mice compared to 15 week-old mature mice. Consistent with this, *in situ* hybridization revealed high level of *Bap1* mRNA expression in terminal end buds (Figure 5c), and higher levels of *Bap1* expression in the ductal epithelium of adolescent as compared with mature female mice. Like *Brca1*, *Bap1* mRNA was expressed in the mammary epithelium at levels higher than those found in the stromal compartment. Steady-state levels of *Bap1* mRNA were up-regulated in the mammary glands of pregnant mice (Figure 5b and c). Like *Brca1*, *in situ* hybridization demonstrated that this up-regulation occurred predominantly in developing alveoli (Figure 5c). *Bap1* was expressed at higher levels in the mammary glands of parous animals that had undergone 28 days of post-lactational regression as compared with age-matched virgin controls (cf. D28 regression and 15 week virgin). The observation that *Bap1* expression is up-regulated in the breast during puberty, pregnancy and as a result of parity is similar to that previously made for *Brca1* and suggests the developmental co-expression of these two proteins (Marquis *et al*., 1995). It should be noted, however, that the magnitude of *Bap1* up-regulation at each of these developmental stages was lower than that observed for *Brca1*. The spatial distribution of *Bap1* expression was also investigated in the testis and ovary (Figure 5c). Like *Brca1*, *Bap1* was expressed at high levels in semeniferous tubules, with lower levels of expression observed in the surrounding connective tissue. *Bap1* was also expressed at high levels throughout the ovary.

### *BAP1 augments the growth suppressive activity of BRCA1*

Several studies have shown that BRCA1 can affect the growth characteristics of cells (Holt *et al*., 1996; Rao *et al*., 1996). We determined whether BAP1 itself may affect cell growth or may alter BRCA1-mediated changes in cell growth. *BRCA1* and *BAP1* cDNAs were co-transfected into MCF7 breast cancer cells and analysed for their effect on colony formation (Figure 6). This cell line was chosen for several reasons: First, it previously has been shown that the growth of these cells is inhibited by the overexpression of *BRCA1* (Holt *et al*., 1996). Second, both Northern and RT–PCR analyses showed that BAP1 is expressed in this cell line (data not shown). Finally, RT–PCR/SSCP analysis of the open reading frame of *BAP1* cDNA prepared from this cell line showed no mutations (data not shown).

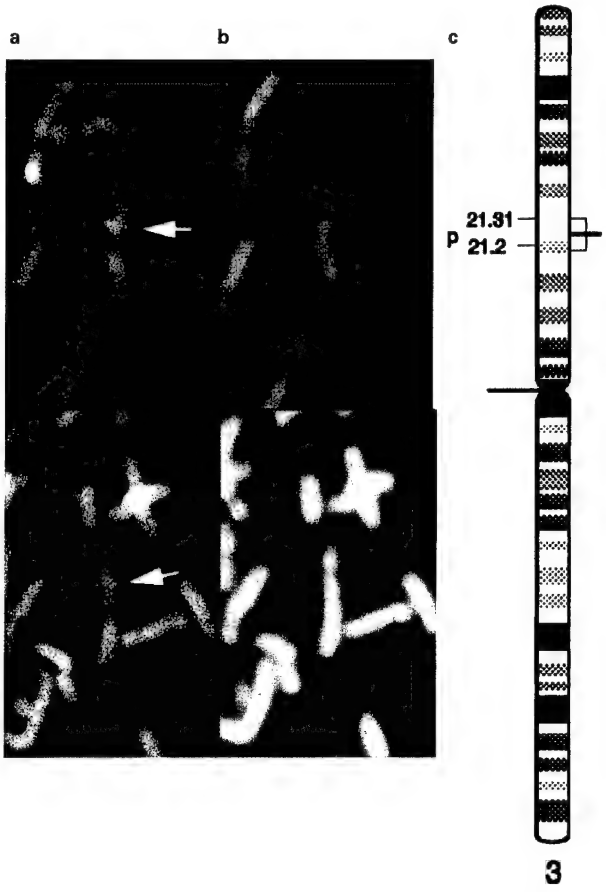



**Figure 7** *BAP1* maps to Chromosome 3p21.3. Fluorescence *in situ* hybridization (FISH) of partial metaphases using biotin-labeled BAP1 cDNA. (**a**) the specific FISH signals on chromosome 3 (arrows), with (**b**) the simultaneously DAPI-stained chromosomes and (**c**) a chromosome ideogram with the localization of BAP1 based on the DAPI-band pattern and FLpter value. The horizontal box indicates the variation in FLpter values on individual chromosomes

The expression of BRCA1 alone (BRCA : pCMV5) decreased the number of colonies formed by these cells when compared to the vector control (pcDNA3 : pCMV5), in agreement with other studies (Holt *et al*., 1996). The co-expression of BRCA1 and BAP1 (BRCA1 : BAP1) significantly decreased the number of cell colonies (approximately fourfold *vs* BRCA1 alone; see Figure 6b) indicating that BAP1 enhances the growth suppressive actions of BRCA1. A mutant of BAP1, BAP1(165–729), in which the enzymatic region is deleted but which still binds to BRCA1, also enhanced the growth suppression of BRCA1, but not to the same extent as the wild-type BAP1.

The expression of BRCA1-Δ11 (a naturally occurring splice variant of BRCA1) in MCF7 cells by itself had no effect on the growth of MCF7 cells (Figure 6).

However, the co-expression of BRCA1-Δ11 and BAP1 dramatically decreased the number of colonies, suggesting that the presence of BAP1 could cooperate with BRCA1-Δ11 in cell growth inhibition. Furthermore, the expression of BAP1 in MCF7 cells also reduced the number of colonies formed (pcDNA3:BAP1; see Figure 6b). The expression of the enzymatic mutant, BAP1(165–729), alone or in combination with BRCA1-Δ11, yielded no growth-suppression activity. Thus, enzymatically active BAP1 enhances BRCA1-mediated suppression of growth in this assay.

### *BAP1 is located on chromosome 3p21.3 and is mutated in non-small cell lung carcinoma*

The data above suggest the possibility that *BAP1* itself may be a tumor suppressor gene and that *BAP1* alterations/deletions might play a role in human tumorigenesis. We determined whether *BAP1* might be located at a chromosomal region previously observed to be mutated in human cancers. The full-length *BAP1* cDNA was used in fluorescent *in situ* hybridization (FISH) to identify the chromosomal location of the *BAP1* gene (Figure 7). Specific signals were observed only on the midportion of the short arm of chromosome 3 with 42 of 69 analysed metaphase spreads showing at least one specific signal. The FLpter value was $0.27 \pm 0.02$ corresponding to a localization for *BAP1* at 3p21.2-p21.31. This location is a region of LOH for breast cancer as well as a region frequently deleted in lung carcinomas (Buchhagen *et al*., 1994; Thiberville *et al*., 1995).

The chromosomal location of *BAP1* suggested the possibility of mutations within *BAP1* in lung tumors. Thus, a total of 44 small cell lung cancer (SCLC), 33 non-small cell lung cancer (NSCLC) and two lymphoblastoid tumor cell lines were screened for mutations within the *BAP1* gene by Southern, Northern and PCR-based SSCP analyses. Genomic DNAs from 15 small cell lung cancer (SCLC), 18 non-small cell lung cancer (NSCLC) and two lymphoblastoid cell lines (representing a subset of the total), were subjected to *Eco*RI digestion and then hybridized to a full-length

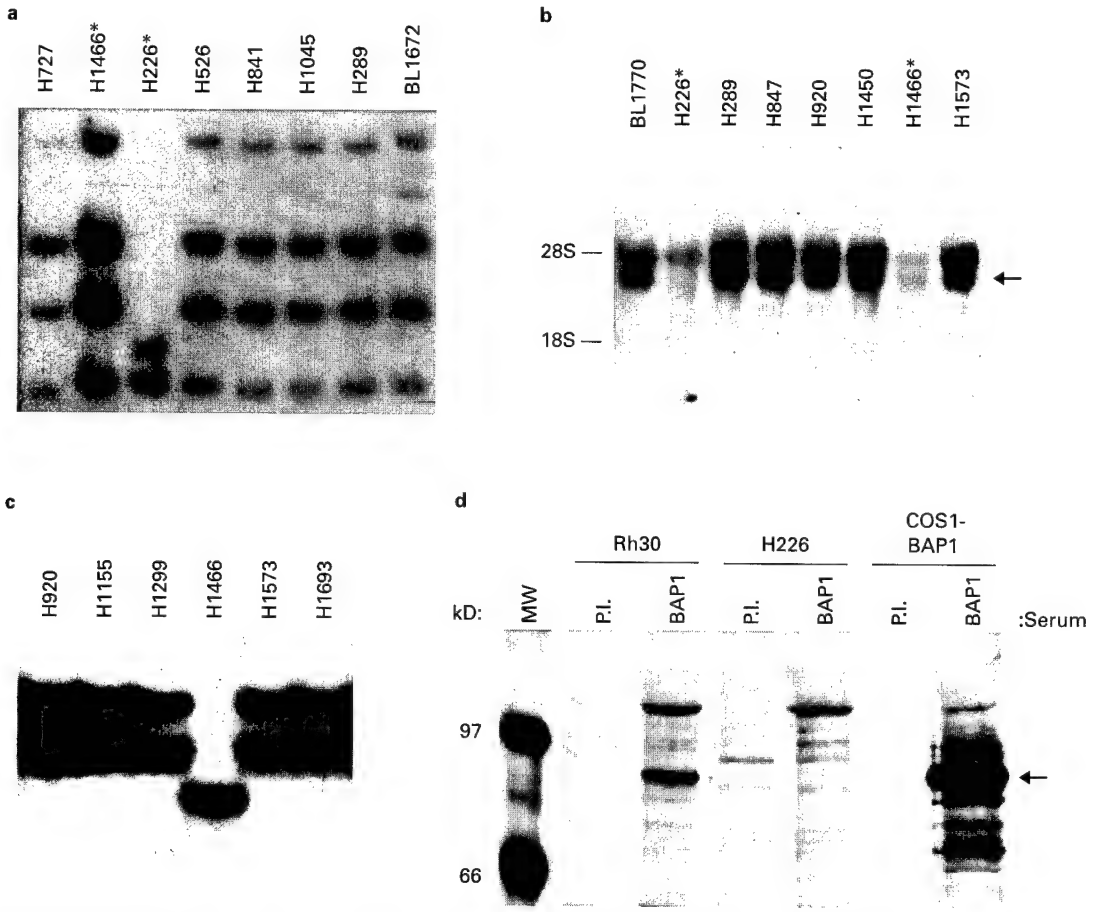



**Figure 8** Mutational analysis of lung carcinomas. **(a)** Southern blot hybridization with *Bam*HI digestion showing four distinct bands at 7.5 kbp, 4.0 kbp, 3.0 kbp and 2.4 kbp as detected by a full-length BAP1 probe. The non-small cell lung cancer NCI-H226 line shows an absence of the 7.5 kbp, 4.0 kbp and 3.0 kbp bands. An aberrant 2.6 kbp band is detected in the NCI-H226 cell line. **(b)** Northern blot hybridization showing a ~4 kb message (arrow) which is absent in the NCI-H226 cell line and the non-small cell lung cancer NCI-H1466 cell line. A fainter (5.0 kb) band is visible corresponding to cross hybridization with the 28S ribosomal component. **(c)** SSCP analysis showing a homozygous shift in H1466 detected by RT–PCR amplification spanning nts 1089 to 1286. **(d)** BAP1 is a 90 kD protein and is missing from NCI-H226 cells. Endogenous BAP1 from Rh30 or H226 cells, or BAP1 expressed in COS1 cells from the BAP1 cDNA, was immunoprecipitated from whole cell lysates prepared after $^{35}$S-labeling of $1 \times 10^7$ cells. Immunoprecipitated proteins were analysed by SDS–PAGE followed by fluorography. The arrow indicates the BAP1 protein. All cell lines in panels **a**, **b** and **c** are listed by their NCI number (Phelps *et al*., 1996)

BAP1 cDNA probe. A single 23 kb band was detected in the lymphoblastoid and most tumor cell lines (data not shown). One NSCLC line, NCI-H226, did not show the 23 kbp band but did show an aberrant 30 kbp band (data not shown). This abnormality was confirmed by *Bam*HI, *Xba*I, *Pst*I and *Bgl*II digestions. Using the 3.5 kbp BAP1 cDNA probe with *Bam*HI digested genomic DNAs, we detected four distinct bands at 7.5 kbp, 4.0 kbp, 3.0 kbp and 2.4 kbp which were present in all cell lines tested with the exception of NCI–H226 (a representative subset is shown in Figure 8a). In the NCI-H226 line, we detected only the 2.4 kbp band and an aberrant 2.6 kbp band.

Further mutational analysis of BAP1 was performed by screening a subset of 31 SCLC and 27 NSCLC lung cancer cell lines and two lymphoblastoid cell lines for expression of *BAP1* mRNA. Northern analysis showed that most cell lines expressed a single ~4 kb mRNA (Figure 8b). However, two cell lines, NCI–H226 and NCI–H1466 (both NSCLCs), showed undetectable levels of *BAP1* expression, suggesting that *BAP1* may be a target for inactivation or down-regulation during NSCLC pathogenesis. The absence of BAP1 protein in the NCI–H226 line was determined by immunoprecipitation of BAP1 from $^{35}$S-labeled cells (Figure 8d).

We screened cDNAs from all 44 SCLC and 33 NSCLC cell lines for mutations in the BAP1 open reading frame by RTPCR–SSCP (Figure 8c). A homozygous, eight base pair deletion and a presumed splice variant were detected. An 8 bp deletion was detected in the cDNA from the NSCLC line NCI–H1466; This short deletion leads to a frameshift/truncation yielding a predicted 393 amino acid protein. This homozygous deletion was confirmed in genomic DNA from the same cell line (data not shown). A 54 bp in-frame deletion (nts. 705–758) was detected in the NCI–H1963 (SCLC) cell line. This deletion was heterozygous at the cDNA level and not present in the genomic DNA, suggesting that it is a splice variant. However, this variant was not detected in any of the other cell lines screened. These data clearly show that genetic alterations, including intragenic homozygous deletions, can occur in *BAP1*.

## Discussion

We have discovered and characterized a novel protein, BAP1, which binds to the BRCA1 RING finger motif. Several lines of evidence are offered which support a role for BAP1 in BRCA1 signal transduction pathways. We showed that: (1) BAP1 binds to the RING finger of BRCA1; but not to germline mutants of BRCA1 or related RING domains; (2) The BAP1–BRCA1 interactions occurs *in vitro* and *in vivo*, and these proteins are partially co-localized in nuclear dot structures; (3) *BAP1* mRNA is expressed in those tissues which also expresses *BRCA1*, and the spatial/temporal distribution of *Bap1* expression in the mouse breast is very similar to that observed with *Brca1*; (4) BAP1 enhances BRCA1-mediated suppression of cell growth in colony formation assays, and suppression by BAP1 is augmented by its UCH enzymatic domain; and (5) *BAP1* maps to chromosome 3p21.3 and is homozygously deleted in a lung carcinoma cell line. Together, these observations suggest that *BAP1* may be a tumor suppressor gene, and that it may serve as a regulator of (or is an effector for) BRCA1 growth control/differentiation pathways. The specificity of the BRCA1 RING finger-BAP1 interaction and the fact that independent, germline missense mutations in the BRCA1 RING finger domain abolish interaction with BAP1 provide substantial evidence for the physiological relevance of this interaction.

BAP1 is a nuclear-localized, ubiquitin carboxy-terminal hydrolase (UCH) which can cleave model ubiquitin substrates *in vitro*. The UCH homology of BAP1 implies a role for either ubiquitin-mediated, proteasome-dependent degradation or other ubiquitin-mediated regulatory (Isaksson *et al*., 1996) pathways in BRCA1 function. Regulated ubiquitination of proteins and subsequent proteasome-dependent proteolysis plays a role in almost every cellular growth, differentiation and homeostatic process (reviewed by Ciechanover, 1994; Isaksson *et al*., 1996; Wilkinson, 1995). The pathway is regulated both at the level of substrate specificity – via the concerted actions of activating enzymes, carrier proteins and ligation enzymes – and at the level of proteolytic deuibiquitation and ubiquitin hydrolysis. The latter enzymes are ubiquitin-specific thiol proteases which have been broadly classified into two families: the ubiquitin-specific protease (UBPs) and the ubiquitin carboxy-terminal hydrolases (UCHs).

The UBP family members are 50–300 kDa, cytoplasmic or nuclear-localized proteins which, in general, cleave ubiquitin or ubiquitin-conjugates from large substrates. Such enzymatic activity can be found directly associated with the 26S proteasome and may serve a regulatory function by editing ubiquitin on large substrates or cleaving polyubiquitin, thus replenishing ubiquitin pools (Lam *et al*., 1997). Remarkably, a number of UBPs have been isolated as growth regulatory and/or developmental control genes such as DOA4 in yeast, which controls DNA replication and repair (Papa and Hochstrasser, 1993); UBP3 which is involved in transcriptional silencing in yeast (Moazed and Johnson, 1996); the TRE2 oncogene which is mutated in the UBP active site and functions as a dominant negative transforming gene (Nakamura *et al*., 1992); the drosophila *Fat Facets* gene which controls pattern formation and eye development (Huang *et al*., 1995; Huang and Fischer-Vize, 1996); and the human DUB family of cytokine-inducible UBPs which control hematopoietic differentiation (Zhu *et al*., 1996, 1997).

By contrast, the UCH family has been characterized as a set of small (25–30 kDa) cytoplasmic proteins which prefer to cleave ubiquitin from ubiquitin-conjugated small substrates and may also be involved in the co-translational processing of proubiquitin. Like the UBPs, UCHs show considerable tissue specificity and developmentally-timed regulation (Wilkinson *et al*., 1992). UCH family members are differentially expressed in neuronal, hematopoietic and germ cells in many species. Most remarkably, a novel UCH enzyme has recently been cloned from *A. californica* whose enzymatic function is essential for acquisition and maintenance of long-term memory (Hedge *et al*., 1997). Finally, UCH levels are down-regulated during viral transformation of fibroblasts (Honore *et al*., 1991), consistent with a role in growth control.

BAP1 is the newest member of the UCH family and considerably expands the potential roles of this family of proteases. BAP1 is a much larger protein (90 kDa) and is the first nuclear-localized UCH. In addition to containing the ~250 amino acid amino-terminal UCH catalytic domain, it includes a carboxy-terminal extension rich in proline, serine and threonine, and a short, highly acidic region; these elements may confer a short half-life upon the protein (Rechsteiner and Rogers, 1996). The extreme carboxy-terminus encodes two potential nuclear localization signals that overlap the approximately 125 amino acid BRCA1-interaction domain. This domain is predicted to fold into a long amphiphatic helix of coiled-coil character, the structure of which may be important for BRCA1 interaction. Indeed, truncation into this region or substitution of a proline for leucine (L691P) abolish the BAP1–BRCA1 interaction in the two-hybrid assay. We have also detected a potential splice variant in BAP1 that results in loss of 31 amino acids of the BRCA1 interaction domain and greatly reduces the ability of BAP1 to bind the BRCA1 RING finger. Thus, our data suggest that the BAP1 carboxy-terminus is tethered to the BRCA1 RING finger domain, leaving the UCH catalytic domain free to interact with ubiquitinated, or other ubiquitin-like, substrates.

A simple model explaining most of our data is that BRCA1 is a direct substrate for the UCH activity of BAP1 and deubiquitination results in the stabilization of BRCA1. Thus, in contrast to the known UCHs which are comprised entirely of the UCH domain, the carboxy-terminal extension of BAP1 may provide substrate and/or targeting specificity for the catalytic function. Paradigms for separate substrate recognition and catalytic domains occur throughout the ubiquitin conjugation/ligation system (see Wilkinson, 1995 and references therein). Regulated ubiquitination of BRCA1 and subsequent proteasome-mediated degradation would be consistent with tight regulation of BRCA1 levels and its subnuclear localization during both the mitotic cell-cycle and meiosis (Gudas *et al*., 1996; Scully *et al*., 1997b; Zabludoff *et al*., 1996). Thus, BAP1-mediated deubiquitination of BRCA1 would stabilize the protein and protect it from proteasome-mediated degradation. This scenario is consistent with both the ability of co-transfected BAP1 to enhance the tumor suppressor effects of BRCA1 in colony formation assays and the finding of mutations in BAP1 in cancer cell lines.

A second, and equally plausible hypothesis is that the BRCA1–BAP1 association serves to target the UCH domain to other substrates that may be bound to other sites on BRCA1. In this scenario, BRCA1 may function as an assembly or scaffold molecule for regulated assembly of multiprotein complexes; a function that has been postulated for other tumor suppressor proteins (e.g. pRb; Sellers and Kaelin, 1996; Welch and Wang, 1995). Thus, BAP1 could be a regulator of this assembly via its control of ubiquitin-mediated proteolysis. Recently, it has been shown that the RING finger protein Ste5 (*S. cerevisiae*) functions as a scaffold protein for assembling protein kinase-dependent signaling complexes in pheromone signaling; this activity is abolished by mutations in the Ste5 RING finger (Inouye *et al*., 1997). In this context, two other RING finger-containing proteins are involved in complexes whereby controlled proteolytic processes are dependent upon the integrity of the RING finger structure: (1) The murine homologue of the drosophila *seven-in-absentia*, siah, (a RING finger protein) binds to the tumor suppressor protein Deleted in Colon Cancer (DCC) and targets it for proteasome-mediated degradation. This degradation requires the siah RING finger structure (Hu *et al*., 1997); and (2) The herpesvirus protein VMW110 is a RING finger protein that binds directly to a UBP family member, HAUSP, and appears to target it to the ND10/POD nuclear dot structure. The ND10 structure itself contains another RING finger protein, the proto-oncogene PML (Everett *et al*., 1997). Remarkably, the PML RING finger has been shown to bind and co-localize with a ubiquitin-like molecule PIC1/SUMO1 which is emerging as a central molecule in nuclear-localized ubiquitin-dependent regulation (Saitoh *et al*., 1997).

A third hypothesis is that BAP1 is involved in the regulation of protein subcellular localization. Mono-ubiquitination, or addition of a ubiquitin-like moiety, has emerged as an important post-translational modification which may affect the specific targeting of proteins to locations other than the proteasome. For example, the addition of the ubiquitin-like PIC1/SUMO1 protein appears to mediate the movement of RanGAP1 (a GTPase-activating protein) from the cytoplasm to the nuclear envelope where it binds the RanBP2 protein (Mahajan *et al*., 1997). This interaction requires the ATP-dependent, covalent addition of PIC1/SUMO-1 to RanGAP1, a process extremely similar to the ATP-dependent, ubiquitin-ligation mechanism. Thus, BAP1-mediated removal of ubiquitin, or ubiquitin-like molecules, from BRCA1, or a protein associated with BRCA1, could target the complex to another cellular compartment, thus altering its function without physically destroying it. That BRCA1 is targeted to specific subcellular sites is evidenced by the observation that it accumulates in nuclear dot structures during S phase of the cell cycle. This localization is abolished at the S/G2 boundary (Scully *et al*., 1997b).

The association of BRCA1 with RAD51 in both mitotic nuclear dot structures and meiotic cells broadly implicates BRCA1 in DNA repair and/or recombination processes. The RAD51/52-dependent DNA repair pathway is highly regulated and includes many proteins, some of which may be potential substrates for BAP1-mediated ubiquitin hydrolysis. RAD23, which associates with the RAD51/52 complex, contains an amino-terminal ubiquitin-like domain which is required for RAD23 function and double-strand break repair (Watkins *et al*., 1993). Recently, a human ubiquitin-like protein, UBL-1, was isolated as a protein which binds directly to the human RAD51/RAD52 complex (Shen *et al*., 1996b). Interestingly, the yeast homologue of UBL1 is SMT3, which functionally associates with the yeast centromere protein MIF2, a protein required for proper chromosome segregation (Brown, 1995; Brown *et al*., 1993). Furthermore, the RAD51/52 complex contains a ubiquitin conjugating enzyme, hUBC9/UBE2I (Jiang and Koltin, 1996; Shen *et al*., 1996a). Thus, it appears that the DNA repair machinery contains both ubiquitin-conjugating and -hydrolyzing

elements, since BAP1 is now implicated as a member of the BRCA1/RAD51/hUBC9 complex. It is possible that BAP1, which is co-expressed with BRCA1 in breast tissue, may regulate the recombination/repair functions of the BRCA1/RAD52 complex by targeting either RAD23 or UBL1 for ubiquitin hydrolysis.

The implications that BAP1 is a key regulator and/or effector of BRCA1 suggest that BAP1 may also play a role in human cancer. The finding that *BAP1* maps to human chromosome 3p21.3 strongly suggested this link; Loss of chromosome 3p genes is a critical event in lung cancer pathogenesis and other carcinomas. Interestingly, two other components of the ubiquitin metabolism pathway have been mapped to chromosome 3p21.3; an ubiquitin activating enzyme (Kok *et al*., 1993) and a ubiquitin protease, UNP (Gray *et al*., 1995). The identification of BAP1 as a UCH suggests a cluster of metabolically related enzymes at this locus. Furthermore, the frequent loss of this chromosomal region suggests that there may be a selective advantage for the loss of ubiquitin-mediated cellular processes during carcinogenesis. That BAP1 may be included in this paradigm is suggested by our detection of rearrangements/deletions within BAP1 in lung cancers (this report) and our detection of independent, homozygous, point mutations in highly conserved residues of BAP1's enzymatic domain (Proctor *et al*., manuscript in preparation).

In summary, we have isolated a novel ubiquitin hydrolase which associates directly with the BRCA1 RING finger domain. BAP1 may play a key role in ubiquitin-dependent regulatory processes in the nucleus including transcription, chromatin remodeling, cell cycle control and DNA repair/recombination.

## Materials and methods

### *Cell culture, transfections and colony formation assays*

All cells were grown at 37°C and 5% $CO_2$. COS1 and MCF7 cells were maintained in DMEM supplemented with 10% fetal bovine serum (FBS), L-glutamine and non-essential amino acids. Small cell lung cancer (SCLC) and non-small cell lung cancer (NSCLC) cell lines were maintained in RPMI media with 10% fetal bovine serum (Gibco BRL). Rh30 cells were maintained in RPMI supplemented with 10% FBS and non-essential amino acids. COS1 cells were transfected using DOSPOR transfection reagent (Boehringer Mannheim Biochemicals) following the manufacturer's protocol.

### *Colony formation assay*

MCF7 cells were transfected by a modified $CaPO_4$-DNA precipitation method (Holt *et al*., 1996). MCF7 cells, at $2 \times 10^6$ cells/10 cm dish, were fed fresh medium approximately 3 h prior to transfection, and were then treated with the Ca-DNA precipitate for 4 h. The cells were subjected to a brief shock with transfection buffer containing 15% glycerol. Twelve to 16 h later, the cells were trypsinized, counted and plated directly into complete medium containing 0.75 mg/mL G418 at $5 \times 10^5$ cells per 10 cm dish. Cells were fed fresh medium containing G418 every 3–4 days. Cells were stained for colonies approximately 21–28 days after transfection.

### *Yeast 2-hybrid*

The yeast 2-hybrid system as modified by Stan Hollenberg was used for all yeast experiments (Vojtek *et al*., 1993). Two libraries were screened for interaction with LexA-BRCA1; a human B cell, oligo-dT-primed, cDNA library (Durfee *et al*., 1993); a kind gift from Dr Steve Elledge) and a mouse embryo (9.5–10.5 day), random-primed, cDNA library size selected for inserts of 300–500 base pairs in length (Vojtek *et al*., 1993); a kind gift of Dr Stan Hollenberg).

### *Construction of expression plasmids*

*LexA fusion constructs:* The 100 amino acid terminal region of human BRCA1 (BRCA1-RF) was used as the bait to screen for interacting proteins. All LexA fusion constructs were made by cloning the appropriate RING (or other) domain into the vector pBTM-116 (Vojtek *et al*., 1993). A synthetic gene of the BRCA1–RF domain was made from overlapping oligonucleotides whose codon usage had been optimized for expression in *E. coli* and *S cerevisiae* (Madden *et al*., 1991). Double-stranded DNA was generated by the Polymerase Chain Reaction (PCR) and amplified with flanking primers containing *Eco*RI and *Sal*I enzymatic restriction sites. A 'wild type' BRCA1–RF domain was confirmed by DNA sequencing. The single amino acid substitutions in the BRCA1–RF domain, BRCA1(C64G) (Cys 64 to Gly), BRCA1(C61G) (Cys 61 to Gly), BRCA1(M18T) (Met18Thr) and BRCA1(R71G) (Arg71 to Gly) as well as the BRCA1–delAG185 truncation mutant were created by PCR-mutagenesis (Ho *et al*., 1989). The BRCA1-del31 truncation mutant was a mis-primed PCR reaction of the BRCA1–RF which was identified by DNA sequencing. The LexA-RPT-1 protein (amino acids 1–100) was made by PCR-mediated amplification of the corresponding nucleotides of a RPT-1 PCR sample (Patarca *et al*., 1988); kindly provided by Dr Harvey Cantor). All clones were confirmed by sequencing. Expression of all constructs in yeast was confirmed by Western analysis using antibodies against the LexA DNA-binding domain (data not shown).

*BAP1 constructs:* A full-length BAP1 cDNA was assembled through the fusion of two overlapping EST clones (the IMAGE Consortium (LLNL); cDNA Clones #46154 and #40642; (Lennon *et al*., 1996)) and the insertion of 62 nucleotides missing from clone #40642 (as revealed by sequencing and RT–PCR analyses). GST-hBAP1(483–729) was generated by cloning the *Xho*I fragment of pAct (the original human two-hybrid clone; nucleotides 1486–3525) into pGEX-5x-1 (Pharmacia Biotech, Inc.). GST-hBAP1(438–594) and pACT-hBAP1(438–594) (nucleotides 1486–1821) were generated and amplified by PCR, digested with restriction enzyme and ligated into the appropriate vector.

### *Mapping of BRCA1/BAP1 interaction domain*

Truncations of mBAP(596–721), and the point mutation mBAP1(L691P), were generated by PCR-based mutagenesis. Products were then ligated into the mouse library-yeast expression vector, pVP16. All clones were confirmed by sequencing and expression in yeast was confirmed by Western analysis using antibodies against the VP16 activation domain (data not shown).

### *Northerns and* in situ

Tissue RNA blots were obtained from Clontech Laboratories, Inc. (Palo Alto, CA). Blots were hybridized with $^{32}$P-labeled hBAP1(483–729) cDNA (nucleotides 1486–3525) using standard protocols. RNA

preparation, Northern hybridization of mouse mammary tissue, and *in situ* hybridization analyses were performed as previously described using the mouse Bap1 cDNA probe described by mBAP1(596–721), corresponding to nucleotides 1825–2202 of the human probe (Marquis *et al*., 1995).

*Fluorescent* in situ *hybridization (FISH)*

FISH using a biotin-labeled 3.5 kb cDNA (full length) clone of BAP1, with corresponding DAPI-banding and measurement of the relative distance from the short arm telomere to the signals (FLpter value) was performed as described previously (Tommerup and Vissing, 1995).

*Immunolocalization*

All immunofluorescence was performed as previously described (Ishov and Maul, 1996). BAP1 protein was detected with affinity purified, BAP1-specific polyclonal antibodies, and BRCA1 was detected with the BR64 monoclonal antibody (Upstate Biotechnology), which were detected with FITC and Texas Red (respectively) using biotin-avidin enhancement. Cells were stained for DNA with bis-benzimide (Hoescht 33258, Sigma Chemical Co.) and mounted using Fluoromount G (Fisher Scientific). Analysis was performed with a confocal scanning microscope (Leica, Inc.).

*BAP1 protein characterization*

*Generation of antibodies:* Using PCR cloning, the cDNA region encoding amino acids 483–576 of BAP1 was fused downstream of the six Histidine residues of the vector pQE-30 (QIAGEN Inc.). The His-tagged protein was purified from *E. coli* over a Ni-agarose column as previously described (Friedman *et al*., 1996) and was used to immunize rabbits for the production of polyclonal antibodies (Cocalico Biologicals, Inc.). Immunoprecipitation of BAP1 was performed by previously described procedures for the metabolic labeling and immunoprecipitation of proteins from cell lysates (Friedman *et al*., 1996).

In vitro *protein association:* GST, GST-hBAP1(483–729) and GST-hBAP1(483–594) were expressed in *E. coli* and then purified as described (Frangioni and Neel, 1993). The $^{35}$S-BRCA1 protein was produced *in vitro* via coupled transcription/translation (TNT®, Promega Corp.). Association between proteins was assayed as described previously (Barlev *et al*., 1995).

*BAP1 enzymatic assay*

Assays for BAP1 enzymatic activity were performed essentially as described for the UCH-L1 and UCH-L3 enzymes (Mayer and Wilkinson, 1989). Briefly, bacteria harboring an IPTG-inducible expression plasmid containing BAP1 (pQE-30; QIAGEN Inc.) were grown and induced with 1 mM IPTG for 4 h. The bacteria were collected and the pellets were resuspended to 1/20 volume (original culture) in lysate buffer (50 mM Tris, pH 8.0, 25 mM EDTA, 10 mM 2-mercapto-ethanol, 100 μg/ml lysozyme). The lysates were sonicated and centrifuged at 40 000 *g*. The soluble fractions were used for subsequent activity assays. The pellets were resuspended in a volume equal to that of the supernatant and samples of both pellet and supernatant were analysed by SDS–PAGE for expression levels and inclusion body formation.

Assays for ubiquitin carboxy-terminal hydrolase activity were performed using the glycine 76 ethyl ester of ubiquitin (Ub-OEt) as substrate (Mayer and Wilkinson, 1989; Wilkinson *et al*., 1986). Assays were done in triplicate. The peak areas were integrated and normalized with respect to a ubiquitin standard.

*Mutation screening*

*RNA/DNA preparation:* Genomic DNA was prepared from breast and lung cancer cell lines using standard methods. Total RNA was extracted by the cesium chloride-ultracentrifugation method (Ausubel *et al*., 1987). First strand cDNAs were synthesized from RNA by M-MLV reverse transcriptase (Gibco–BRL) according to the manufacturer's instructions.

*Southern and Northern blot hybridization:* Five μg of genomic DNA, subjected to restriction enzyme digestion, or 10 μg total RNA, was electrophoretically gel-fractionated and transferred to Hybond N$^+$ membranes (Amersham). Hybridization was performed with a $^{32}$P-full-length BAP1 cDNA probe followed by washes under standard conditions and detection by autoradiography.

*Single strand conformational polymorphism (SSCP) analysis:* Seventeen overlapping PCR primer pairs, each with a predicted product size of approximately 200 base pairs, were designed to span the 2.2 kb open reading frame of the BAP1 cDNA sequence. cDNA (from RNA) was amplified in 20 μl PCR reactions containing 20 mM Tris HCl (pH 8.3), 50 mM KCl, 1.5 mM MgCl, 0.2 mM each dNTP, 0.1 mM each forward and reverse primer, 0.05 ml [$^{32}$P-α]dCTP and 0.5 units Taq DNA Polymerase (BRL). PCR reactions were carried out in a Perkin-Elmer 9600 Thermocycler using a touchdown technique: a 2.5 min initial denaturation at 94°C was followed by 35 cycles of denaturation at 94°C × 30 s, annealing, initially at 65°C decreasing by 1°C for each of the first ten cycles to 55°C, × 30 s, and extension at 72°C × 30 s with a final extension of 5 min at 72°C. PCR products were then diluted 1 : 10 with SSCP dye (95% formamide, 20 mM EDTA, and 0.05% each of bromophenol blue and xylene cyanol), heat-denatured, and electrophoresed on 0.5 × MDE gels ± 10% glycerol. Abnormal single stranded DNA detected as autoradiographic shifts were re-amplified by PCR and subjected to automated dye-terminator sequencing (ABI 373).

**Acknowledgements**

We thank E Koonin (NIH/NCBI) for DNA sequence analysis and suggesting a potential gap in our original cDNA clone; Dr R Baer for supplying an anti-BAP1 antibody which allowed us to verify a variety of results; Drs B Weber, F Couch, W Fredericks and J Friedman for their many helpful discussions and the technical assistance of Jing Wang, Bodil Olsen and Winni Pedersen. JM is a recipient of an NCI P50 Lung Cancer SPORE grant. KDW is supported by NIH grant GM30308. NT is supported by The Danish Cancer Society. The Danish Environment Research Programme, The Danish Biotechnological Research and Development Programme, The Danish Research Center for Growth and Regeneration, The Novo Nordisk Foundation, and The Aage Bang Foundation. LAC is a Charles E Culpeper Medical Scholar and is supported by the Charles E Culpeper Foundation and NCI grant CA71513. GCP is supported by an ACS Junior Faculty Award and is a Pew Scholar in the Biomedical Sciences. FJR is supported by National Institutes of Health grants: Core grant CA10815, DK 49210, GM 54220, DAMD17-96-6141, ACS NP-954, the Irving A Hansen Memorial Foundation, the Mary A Rumsey Memorial Foundation and the Pew Scholars Program in the Biomedical Sciences. DEJ is a Susan G Koman Breast Cancer Foundation Postdoctoral Fellow. GenBank accession No. AF045581.

## References

Altschul SF, Gish W, Miller W, Myers EW and Lipman DJ. (1990). *J. Mol. Biol.*, **215,** 403–410.

Ausubel FM, Brent R, Kingston RE, Moore DD, Seidman JG, Smith JA and Struhl K. (1987). *Current Protocols in Molecular Biology*. John Wiley & Sons, Inc., Boston.

Barlev NA, Candau R, Wang L, Darpino P, Silverman N and Berger SL. (1995). *J. Biol. Chem.*, **270,** 19337–19344.

Borden KLB, Boddy MN, Lally J, O'Reilly NJ, Martin S, Howe K, Solomon E and Freemont PS. (1995). *EMBO J.*, **14,** 1532–1541.

Brown MT. (1995). *Gene*, **160,** 111–116.

Brown MT, Goetsch L and Hartwell LH. (1993). *J. Cell Biol.*, **123,** 387–403.

Buchhagen DL, Qiu L and Etkind P. (1994). *Int. J. Cancer*, **57,** 473–479.

Chapman MS and Verma IM. (1996). *Nature*, **382,** 678–679.

Chen CF, Li S, Chen Y, Chen PL, Sharp ZD and Lee WH. (1996a). *J. Biol. Chem.*, **271,** 32863–32868.

Chen Y, Chen CF, Riley DJ, Allred DC, Chen PL, Von Hoff D, Osborne CK and Lee WH. (1995). *Science*, **270,** 789–791.

Chen Y, Farmer AA, Chen CF, Jones DC, Chen PL and Lee WH. (1996b). *Cancer Res.*, **56,** 3168–3172.

Ciechanover A. (1994). *Biological Chemistry Hoppe-Seyler*, **375,** 565–581.

Couch FJ and Weber BL. (1996). *Hum. Mutat.*, **8,** 8–18.

Durfee T, Becherer K, Chen PL, Yeh SH, Yang Y, Kilburn AE, Lee WH and Elledge SJ. (1993). *Genes Dev.*, **7,** 555–569.

Easton DF, Bishop DT, Ford D and Crockford GP. (1993). *Am. J. Hum. Genet.*, **52,** 678–701.

Easton DF, Ford D, Bishop DT and Consortium T.B.C.L. (1995). *Am. J. Hum. Genet.*, **56,** 265–271.

Everett RD, Meredith M, Orr A, Cross A, Kathoria M and Parkinson J. (1997). *EMBO J.*, **16,** 566–577.

FitzGerald MG, MacDonald DJ, Krainer M, Hoover I, O'Neil, Unsal H, Silva-Arrieto S, Finkelstein DM, Beer-Romero P, Englert C, Sgroi DC, Smith BL, Younger JW, Garber JE, Duda RB, Mayzel KA, Isselbacher KJ, Friend SH and Haber DA. (1996). *N. Engl. J. Med.*, **334,** 143–149.

Ford D, Easton DF, Bishop DT, Narod SA and Goldgar DE. (1994). *Lancet*, **343,** 692–695.

Frangioni JV and Neel BG. (1993). *Anal. Biochem.*, **210,** 179–187.

Friedman JR, Fredericks WJ, Jensen DE, Speicher DW, Huang XP, Neilson EG and Rauscher III, FJ. (1996). *Genes & Development*, **10,** 2067–2078.

Futreal PA, Liu Q, Shattuck-Eidens D, Cochran C, Harshman K, Tavtigian S, Bennett LM, Haugen-Strano A, Swensen J, Miki *et al.* (1994). *Science*, **266,** 120–122.

Gray DA, Inazawa J, Gupta K, Wong A, Ueda R and Takahashi T. (1995). *Oncogene*, **10,** 2179–2183.

Gudas JM, Nguyen H, Li T and Cowan KH. (1995). *Cancer Res.*, **55,** 4561–4565.

Gudas JM, Tao L, Nguyen H, Jensen D, Rauscher III FJ and Cowan KH. (1996). *Cell Growth and Differentiation*, **7,** 717–723.

Hakem R, de la Pompa JL, Sirard C, Mo R, Woo M, Hakem A, Wakeham A, Potter J, Reitmair A, Billia F, Firpo E, Hui CC, Roberts J, Rossant J and Mak TW. (1996). *Cell*, **85,** 1009–1023.

Hall JM, Lee MK, Newman B, Morrow JE, Anderson LA, Huey B and King M-C. (1990). *Science*, **250,** 1684–1689.

Hedge AN, Inokuchi K, Pei W, Casadio A, Grirardi M, Chain DG, Martin KC, Kandel ER and Schwartz JH. (1997). *Cell*, **89,** 114–126.

Ho SN, Hunt HD, Horton RM, Pullen JK and Pease LR. (1989). *Gene*, **77,** 51–59.

Holt JT, Thompson ME, Szabo C, Robinson-Benion C, Arteaga CL, King M-C and Jensen RA. (1996). *Nature Genetics*, **12,** 298–302.

Honore B, Rasmussen HH, Vandekerckhove J and Celis JE. (1991). *FEBS Lett.*, **280,** 235–240.

Hu G, Zhang S, Vidal M, La Baer J, Xu T and Fearon ER. (1997). *Genes & Dev.*, **11,** 2701–2714.

Huang Y, Baker RT and Fischer-Vize JA. (1995). *Science*, **270,** 1828–1831.

Huang Y and Fischer-Vize JA. (1996). *Development*, **122,** 3207–3216.

Inouye C, Dhillon N and Thorner J. (1997). *Science*, **278,** 103–106.

Isaksson A, Musti AM and Bohmann D. (1996). *Biochimica et Biophysica Acta*, **1288,** F21–F29.

Ishov AM and Maul GG. (1996). *J. Cell Biol.*, **134,** 815–826.

Jiang W and Koltin Y. (1996). *Mol. Gen. Genet.*, **251,** 153–160.

Jin Y, Xu XL, Yang M-CW, Wei F, Ayi T-C, Bowcock AM and Baer R. (1997). *Proc. Natl. Acad. Sci. USA*, **94,** 12075–12080.

Johnston SC, Larsen CN, Cook WJ, Wilkinson KD and Hill CP. (1997). *EMBO J.*, **16,** 3787–3796.

Klug A and Schwabe JW. (1995). *FASEB J.*, **9,** 597–604.

Kok K, Hofstra R, Pilz A, van den Berg A, Terpstra P, Buys CH and Caritt B. (1993). *Proc. Natl. Acad. Sci. USA*, **90,** 6071–6075.

Koonin EV, Altschul SF and Bork P. (1996). *Nature Genet.*, **13,** 266–268.

Lam YA, Xu W, DeMartino GN and Cohen RE. (1997). *Nature*, **385,** 737–740.

Larsen CN, Price JS and Wilkinson KD. (1996). *Biochem.*, **35,** 6735–6744.

Lennon GG, Auffray C, Polymeropoulos M and Soares MB. (1996). *Genomics*, **33,** 151–152.

Liu CY, Flesken-Nikitin A, Li S, Zeng Y and Lee WH. (1996). *Genes & Dev.*, **10,** 1835–1843.

Lovering R, Hanson IM, Borden KL, Martin S, O'Reilly NJ, Evan GI, Rahman D, Pappin DJ, Trowsdale J and Freemont PS. (1993). *Proc. Natl. Acad. Sci. USA*, **90,** 2112–2116.

Madden SL, Cook DM, Morris JF, Gashler A, Sukhatme VP and Rauscher III, FJ. (1991). *Science*, **253,** 1550–1553.

Mahajan R, Delphin C, Guan T, Gerace L and Melchoir F. (1997). *Cell*, **88,** 97–107.

Marks JR, Huper G, Vaughn JP, Davis PL, Norris J, McDonnell DP, Wiseman RW, Futreal PA and Iglehart JD. (1997). *Oncogene*, **14,** 115–121.

Marquis ST, Rajan JV, Wynshaw-Boris A, Xu J, Yin GY, Abel KJ, Weber BL and Chodosh LA. (1995). *Nature Genet.*, **11,** 17–26.

Mayer AN and Wilkinson KD. (1989). *Biochemistry*, **28,** 166–172.

Merajver SD, Pham TM, Caduff RF, Chen M, Poy EL, Cooney KA, Weber BL, Collins FS, Johnston C and Frank TS. (1995). *Nat. Genet.*, **9,** 439–443.

Miki Y, Swensen J, Shattuck-Eidens D, Futreal PA, Harshman K, Tavtigian S, Liu Q, Cochran C, Bennett LM, Ding W *et al.* (1994). *Science*, **266,** 66–71.

Moazed D and Johnson D. (1996). *Cell*, **86,** 667–677.

Muto MG, Cramer DW, Tangir J, Berkowitz R and Mok S. (1996). *Cancer Res.*, **56,** 1250–1252.

Nakamura T, Hillova J, Mariage-Samson R, Onno M, Huebner K, Cannizzaro LA, Boghosian-Sell L, Croce CM and Hill M. (1992). *Oncogene*, **7,** 733–741.

Papa FR and Hochstrasser M. (1993). *Nature*, **366,** 313–319.

Patarca R, Freeman GJ, Schwartz J, Singh RP, Kong QT, Murphy E, Anderson Y, Sheng FY, Singh P, Johnson KA *et al.* (1988). *Proc. Natl. Acad. Sci. USA*, **85,** 2733–2737.

Phelps RM, Johnson BE, Ihde DC, Gazdar AF, Carbone DP, McClintock PR, Linnoila RI, Matthews MJ, Bunn Jr, PA, Carney D, Minna JD and Mulshine JL. (1996). *J. Cell Biochem. Suppl.*, **24,** 32–91.

Rao VN, Shao N, Ahmak M and Reddy ESP. (1996). *Oncogene,* **12,** 523–528.

Rechsteiner M and Rogers SW. (1996). *Trends Biochem. Sci.,* **21,** 267–271.

Roa BB, Boyd AA, Volcik K and Richards CS. (1996). *Nature Genet.,* **14,** 185–187.

Saitoh H, Pu RT and Dasso M. (1997). *Trends Biochem. Sci.,* **22,** 374–376.

Saurin AJ, Borden KLB, Boddy MN and Freemont PS. (1996). *Trends Biochem. Sci.,* **21,** 208–214.

Scully R, Anderson SF, Chao DM, Wanjiang W, Liyan Y, Young RA, Livingston DM and Parvin JD. (1997a). *Proc. Natl. Acad. Sci. USA,* **94,** 5605–5610.

Scully R, Chen J, Plug A, Xiao Y, Weaver D, Feunteun J, Ashley T and Livingston DM. (1997b). *Cell,* **88,** 265–275.

Scully RSG, Brown M, Caprio JAD, Cannistra SA, Feunteun J, Schnitt S and Livingston DM. (1996). *Science,* **272,** 123–126.

Sellers WR and Kaelin WG. (1996). *Biochim. Biophys. Acta,* **1288,** M1–5.

Shen Z, Pardington-Purtymun PE, Comeaux JC, Moyzis RK and Chen DJ. (1996a). *Genomics,* **37,** 183–186.

Shen Z, Pardington-Purtymun PE, Comeaux JC, Moyzis RK and Chen DJ. (1996b). *Genomics,* **36,** 271–279.

Smith SA, Easton DG, Evans DGR and Ponder BAJ. (1992). *Nature Genetics,* **2,** 128–131.

Struewing JP, Abeliovich D, Peretz T, Avishai N, Kaback MM, Collins FS and Brody LC. (1995). *Nature Genetics,* **11,** 198–200.

Szabo CI and King M-C. (1995). *Hum. Mol. Genet.,* **4,** 1811–1817.

Thakur S, Zhang HB, Peng Y, Le H, Carroll B, Ward T, Yao J, Farid LM, Couch FJ, Wilson RB and Weber BL. (1997). *Mol. Cell. Biol.,* **17,** 444–452.

Thiberville L, Bourguignon J, Metayer J, Bost F, Diarra-Mehrpour M, Bignon J, Lam S, Martin JP and Nouvet G. (1995). *Int. J. Cancer,* **64,** 371–377.

Thompson JD, Higgins DG and Gibson TJ. (1994). *Nucleic Acids Res.,* **22,** 4673–4680.

Thompson ME, Jensen RA, Obermiller PS, Page DL and Holt JT. (1995). *Nat. Genet.,* **9,** 444–450.

Tommerup N and Vissing H. (1995). *Genomics,* **27,** 259–264.

Vaughn JP, Davis PL, Jarboe MD, Huper G, Evans AC, Wiseman RW, Berchuck A, Iglehart JD, Futreal A and Marks JR. (1996). *Cell Growth Differ.,* **7,** 711–715.

Vojtek AB, Hollenberg SM and Cooper JA. (1993). *Cell,* **74,** 205–214.

Watkins JF, Sung P, Prakash L and Prakash S. (1993). *Mol. Cell. Biol.,* **13,** 7757–7765.

Welch PJ and Wang JY. (1995). *Genes Dev.,* **9,** 31–46.

Wilkinson KD. (1995). *Ann. Rev. Nutrition,* **15,** 161–189.

Wilkinson KD, Cox MJ, Mayer AN and Frey T. (1986). *Biochemistry,* **25,** 6644–6649.

Wilkinson KD, Deshpande S and Larsen CN. (1992). *Biochem. Soc. Trans.,* **20,** 631–637.

Wilkinson KD, Lee KM, Deshpande S, Duerksen-Hughes P, Boss JM and Pohl J. (1989). *Science,* **246,** 670–673.

Wu LC, Wang ZW, Tsan JT, Spillman MA, Phung A, Xu XL, Yang MC, Hwang LY, Bowcock AM and Baer R. (1996). *Nature Genet.,* **14,** 430–440.

Zabludoff SD, Wright WW, Harshman K and Wold BJ. (1996). *Oncogene,* **13,** 649–653.

Zhu Y, Carroll M, Papa FR, Hochstrasser M and D'Andrea AD. (1996). *Proc. Natl. Acad. Sci. USA,* **93,** 3275–3279.

Zhu Y, Lambert K, Corless C, Copeland NG, Gilbert DJ, Jenkins NA and D'Andrea AD. (1997). *J. Biol. Chem.,* **272,** 51–57.

[CANCER RESEARCH 59, 5275–5282, October 15, 1999]

# Structure-Function Studies of the BTB/POZ Transcriptional Repression Domain from the Promyelocytic Leukemia Zinc Finger Oncoprotein[1]

**Xinmin Li, Hongzhuang Peng, David C. Schultz, Jesus M. Lopez-Guisa,[2] Frank J. Rauscher III, and Ronen Marmorstein[3]**

*The Wistar Institute [X. L., H. P., D. C. S., J. M. L-G., F. J. R. III, R. M.], Department of Chemistry [R. M.], and the Department of Biochemistry and Biophysics [X. L., R. M.], University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA*

## ABSTRACT

**The evolutionarily conserved BTB/POZ domain from the promyelocytic leukemia zinc finger (PLZF) oncoprotein mediates transcriptional repression through the recruitment of corepressor proteins containing histone deacetylases in acute promyelocytic leukemia. We have determined the 2.0 Å crystal structure of the BTB/POZ domain from PLZF (PLZF-BTB/POZ), and have carried out biochemical analysis of PLZF-BTB/POZ harboring site-directed mutations to probe structure-function relationships. The structure reveals a novel $\alpha/\beta$ homodimeric fold in which dimer interactions occur along two surfaces of the protein subunits. The conservation of BTB/POZ domain residues at the core of the protomers and at the dimer interface implies an analogous fold and dimerization mode for BTB/POZ domains from otherwise functionally unrelated proteins. Unexpectedly, the BTB/POZ domain forms dimer-dimer interactions in the crystals, suggesting a mode for higher-order protein oligomerization for BTB/POZ-mediated transcriptional repression. Biochemical characterization of PLZF-BTB/POZ harboring mutations in conserved residues involved in protein dimerization reveals that the integrity of the dimer interface is exquisitely sensitive to mutation and that dimer formation is required for wild-type levels of transcriptional repression. Interestingly, similar mutational analysis of residues within a pronounced protein cleft along the dimer interface, which had been implicated previously for interaction with corepressors, has negligible effects on dimerization or transcriptional repression. Together, these studies form a structure-function framework for understanding BTB/POZ-mediated oligomerization and transcriptional repression properties.**

## INTRODUCTION

The *PLZF*[4] gene was first identified as part of a t(11;17) chromosomal translocation with the *RARα* gene forming the PLZF-RARα fusion protein in APL (1). Unlike the more common APL t(15;17) translocation forming the PML-RARα fusion protein (2, 3), patients harboring the t(11;17) chromosomal translocation are resistant to treatment with pharmacological doses of RAs (4).

The PLZF moiety of the PLZF-RARα fusion protein contains a 120 residue BTB/POZ domain, named for its presence in the *Drosophila* proteins BTB (5), and its homology with several POZ (6). This domain has been found in an increasing number of proteins in poxvirus, *Caenorhabditis elegans*, *Drosophila*, and humans and is generally found at the $NH_2$ terminus of either actin-binding or, more commonly, nuclear transcriptional regulatory proteins (7). Functional studies have shown that the BTB/POZ domain mediates homodimerization (6, 8, 9), heteromultimerization between different BTB/POZ-harboring proteins (6, 10), and transcriptional repression in the case of several DNA regulators harboring BTB/POZ domains (10–16).

Much of the mechanistic detail for the function of BTB/POZ domains has come from the study of the PLZF and LAZ3/BCL6 (lymphoma-associated zinc finger 3/B cell lymphoma 6) oncoproteins (10, 11). In these cases, the BTB/POZ domain has been shown to promote transcriptional repression through the recruitment of corepressor proteins such as N-CoR and SMRT (17, 18). More recently, the BTB/POZ domain of PLZF has been shown to interact with a protein complex containing N-CoR/SMRT, mSin3A, and the histone deacetylase, HDAC1, to mediate transcriptional repression (19–21). Moreover, this recruitment has been found to play a major role in the pathogenic effect of the PLZF-RARα fusion protein and for its resistance to treatment with RA. Specifically, a model has been proposed whereby the PLZF-RARα fusion protein acts as a potent transcriptional repressor through the ability of both the RARα (22) and BTB/POZ moieties to recruit the SMRT/N-CoR deacetylase transcriptional repression complex (23, 24). Because RA induces the release of this corepressor complex from RARα (22) but not from PLZF, this model is consistent with the RA resistance of APL patients harboring the PLZF-RARα translocation (25, 26).

Here we present the high resolution crystal structure of the BTB/POZ domain from PLZF and characterize the biochemical properties of PLZF proteins harboring site-directed mutations. The structure provides general insights into the architecture and mode of multimerization for the evolutionarily conserved BTB/POZ domain. Moreover, a correlation of the BTB/POZ domain structure with the dimerization and transcriptional repression properties of PLZF proteins harboring site-directed mutations establishes a structure-function paradigm for understanding the dimerization and transcriptional repression properties of proteins harboring BTB/POZ domains. Finally, the insights provided here provide a framework from which to design PLZF-specific inhibitory molecules that may be used to treat APL patients harboring the PLZF-RARα translocation.

## MATERIALS AND METHODS

**Protein Expression and Purification for Crystallization.** Residues 6–123 of PLZF harboring the BTB/POZ domain and containing the $NH_2$-terminal 6xHis tag sequence MRGSHHHHHHGS (herein called PLZF-BTB/POZ) was overexpressed using a pQE30 T5-polymerase based expression vector in *Escherichia coli* S9 cells and purified using a combination of anion-exchange (Q-Sepharose) and gel filtration (Superdex-75) chromatography as described elsewhere (8).

SeMet-derivatized PLZF-BTB/POZ protein was prepared by growing pQE30/PLZF-BTB/POZ-transformed *E. coli* strain B834 (DE3; Novagen) in 4-morpholinepropanesulfonic acid-based minimal media (27) supplemented with 50 mg/l L-SeMet and other amino acids at the suggested concentrations. Cells were grown at 28°C to an $A_{595}$ of 0.4 and induced with 1 mM isoprpyl-1-thio-β-D-galactopyranoside to an $A_{595}$ of ~1.0. The PLZF-BTB/POZ protein was isolated essentially as described for the underivatized protein. Quantitative

Received 6/14/99; accepted 8/19/99.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

[1] Supported by a grant from the Leukemia Research Foundation (to R. M.) and NIH Basic Cancer Research Training Grant CA09171 (to D. C. S). F. J. R. was supported in part by NIH Grant CA52009, Core Grant CA10815, Core Grant DK50306, and Grants DK49210, GM54220, DAMD17-96-1-6141, and ACS NP-954; the Irving A. Hansen Memorial Foundation; the Mary A. Rumsey Memorial Foundation; and the Pew Scholars Program in the Biomedical Sciences.

[2] Present address: Childrens Hospital and Medical Center, Seattle, WA 98105.

[3] To whom requests for reprints should be addressed, at The Wistar Institute, 3601 Spruce Street, Philadelphia, PA 19104. Phone: (215) 898-5006; Fax: (215) 898-0381; E-mail: marmor@wistar.upenn.edu.

[4] The abbreviations used are: PLZF, promyelocytic leukemia zinc finger; RAR, retinoic acid receptor; APL, acute promyelocytic leukemia; RA, retinoic acid; BTB, *B*road complex, *T*rametrack, and *B*ric a brac; POZ, *po*xvirus and *z*inc finger proteins; SMRT, silencing mediator of retinoid and thyroid receptor; SeMet, selenomethionine; MAD, multiwavelength anomalous dispersion.

amino acid analysis of SeMet-derivatized PLZF-BTB/POZ protein confirmed that >90% of the methionine residues had been replaced. After purification, SeMet-derivatized PLZF-BTB/POZ was concentrated to ~50 mg/ml by centrifugation using a Centricon-10 microconcentrator (Amicon) in a buffer containing 40 mM Tris (pH 8.5), 100 mM NaCl, and 1 mM β-mercaptoethanol and frozen as 50-μl aliquots at −70°C before crystallization. Frozen protein aliquots were thawed for use in crystallization as needed.

**Crystallization and Data Collection.** Crystals of underivatized and SeMet-derivatized PLZF-BTB/POZ were prepared using 2-μl hanging drops containing 10 mg/ml PLZF-BTB/POZ, 8% isopropanol, 600 mM $MgCl_2$, 50 mM Tris (pH 8.5), and 50 mM HEPES (pH 6.5) equilibrated over a reservoir containing two times the concentration of salts, buffer, and precipitating agent. Crystals were transiently transferred (for ~5 min) to a harvest solution composed of salts, buffer, and precipitating agent at the same concentrations as the reservoir solution with the addition of 25% glycerol to facilitate X-ray data collection at cryogenic temperature (−170°C).

MAD data were collected from cryoprotected SeMet-derivatized PLZF-BTB/POZ crystals that were flash frozen in liquid propane and stored in a Dewar-containing liquid nitrogen prior to data collection at 110 K. MAD data were collected at NSLS using beamline X4A equipped with an R-AXIS IV image plate detector. The inverse-beam method was used to record Bijvoet differences from each of four different wavelengths to optimize dispersive differences: upstream remote ($\lambda_1$= 0.9878 Å), the downstream remote ($\lambda_4$= 0.9667 Å), the inflection ($\lambda_2$= 0.9796 Å), and the maximum of X-ray absorption ($\lambda_3$= 0.9795 Å). The MAD data were processed with DENZO and SCALEPACK (Ref. 28; Table 1).

**Structure Determination and Refinement.** The structure was determined to 2.3 Å resolution using the MAD data collected from SeMet-derivatized PLZF-BTB/POZ crystals. SeMet positions were identified and refined using difference Patterson and difference Fourier synthesis with the PHASES package (Ref. 29; Table 1). Two SeMet sites were unambiguously identified using difference Patterson maps produced with two sets of dispersive signals and three sets of Bijvoet signals, and two additional SeMet sites were confirmed using cross-difference Fourier maps using the first two sites. Phase refinement, using four SeMet sites using all Bijvoet and dispersive differences, and solvent flattening procedures were carried out with the program PHASES (29). The resulting electron density map was of very high quality (Fig. 1*C*) and allowed for straightforward tracing of the polypeptide chain with the program O (30) using the SeMet positions as landmarks.

Model refinement was carried out with the program X-PLOR (31) using data collected from SeMet-derivatized PLZF-BTB/POZ crystals at the upstream remote wavelength (0.9878 Å). Before refinement, a randomly selected set of data (10%) was omitted from the refinement and used as a "free data set" to monitor subsequent calculations (32), and the model was refined against the remaining data (90%, working data set). Conventional position refinement was initially carried out at 3.0 Å with X-PLOR. After this procedure, iterative cycles of positional refinement with X-PLOR (33) and manual model building with O (30) using Sigmaa-weighted 2Fo-Fc and Fo-Fc difference maps were extended in steps to resolution limits of 2.7, 2.5, 2.3, and 2.0 Å. The later stages of refinement using both simulated annealing (34) and torsion angle dynamics (35) as implemented in X-PLOR. Also at the later stages of refinement, a bulk solvent correction was applied (36), tightly constrained atomic B-factors were adjusted, and water molecules were built into regions that showed strong Fo-Fc peaks and made stereochemically feasible hydrogen bonds. The correctness of the model was checked against simulated annealing omit maps (37) over the entire structure by omitting 15 residues at a time, and the model was adjusted appropriately. A last round of refinement resulted in a model with good geometry ($RMS_{bond\ length}$ = 0.005 Å, $RMS_{bond\ angle}$ = 0.862°) and a working R factor of 25.2% with a free R factor of 27.3% using all reflections between 20 and 2 Å (Table 1). The final model includes residues 7–122 of PLZF-BTB/POZ and 61 water molecules. Residue 6, the $NH_2$-terminal 6xHis-tag and the COOH-terminal residue are not visible in the electron density map. A Ramachandran plot showed no residues of the protein in disallowed regions (38).

**Site-directed Mutagenesis.** The plasmids containing the BTB/POZ domain of PLZF (residues 6–123) were constructed by PCR using the plasmid pQE30-PLZF as a template (8). A 5′ oligonucleotide (5′-GGA TCC ACC ATG GGC ATG ATC CAG CTG CAG-3′) with a *Bam*HI site immediately 5′ to a consensus Kozak sequence (ACC) at methionine 6 and a 3′ oligonucleotide (5′-GAT GGA TCC CTA CTC CAG CAT CTT CAG GCA CTG-3′) with a stop codon (TAG) and *Bam*HI site after amino acid 123 were used to amplify the desired sequence. Single amino acid point mutations within the BTB/POZ domain of $PLZF_{6-123}$ were created using standard PCR-mediated mutagenesis. The mutagenic primers contained the following codons: L21A, CTG to GCG; D35N, GAT to AAT; H64A, CAC to GCC; N66A, AAT to GCT; and Q68A, CAA to GCA. *Bam*HI-digested PCR products were ligated into *Bam*HI-digested pSP73 for *in vitro* translation and pM2 for *in vivo* expression. All PCR-derived plasmids were subjected to automated DNA sequencing of both strands to confirm the incorporation of appropriate mutations and integrity of surrounding sequences.

**Gel Filtration Analysis of Wild-Type and Mutant PLZF Proteins.** Fifty μl of [$^{35}$S]methionine-labeled, *in vitro*-translated $PLZF_{6-123}$ proteins (SP6 TnT; Promega) were analyzed by gel filtration with a Superdex 200 HR 10/30 column (Pharmacia Biotech, Inc.) equilibrated in PBS (10 mM $Na_2HPO_4$, 1.4 mM $KH_2PO_4$, 137 mM NaCl, and 2.7 mM KCl, pH 7.0). The column was run at 4°C at a flow rate of 0.3 ml/min, and 1-ml fractions were collected. The

Table 1 *Crystallographic data*

Unit cell parameters

PLZF-BTB/POZ (residue 7-122; four ordered SeMet)
Space group I222 a = 38.8 Å b = 78.2 Å c = 85.2 Å

Data collection (between 20 and 2 Å)

| Wavelength (Å) | Unique reflections | Total reflections | Completeness | $R_{merge}$[a] | Ave I/σ |
|---|---|---|---|---|---|
| $\lambda_1$ 0.9879 | 8,790 | 115110 | 96.7% | 7.0% | 35.5 |
| $\lambda_2$ 0.9795 | 16,465 | 115154 | 97.3% | 7.2% | 27.0 |
| $\lambda_3$ 0.9791 | 16,460 | 116531 | 97.3% | 7.8% | 26.8 |
| $\lambda_4$ 0.9667 | 15,963 | 109193 | 94.3% | 8.0% | 25.5 |

Phasing statistics[b] (between 20 and 2 Å)

| Difference (w.r.t $\lambda_1$) | Phasing power[d] | FOM Centric | FOM Accentric |
|---|---|---|---|
| $\lambda_2$ disp | 2.43 | 0.541 | 0.368 |
| $\lambda_2$ anom | 2.63 | 0.335 | |
| $\lambda_3$ disp | 1.67 | 0.445 | 0.284 |
| $\lambda_3$ anom | 3.35 | 0.384 | |
| $\lambda_4$ anom | 2.16 | 0.292 | |
| Combined FOM | | 0.694 | |

Refinement statistics[c]

| | | |
|---|---|---|
| Resolution range | | Between 20 and 2 Å |
| Number of refl. | | 7889 ($I/\sigma>2$)[f] |
| $R_{working}/R_{free}$[e] (%) | | 25.2/27.3 |
| RMS deviations | Bonds | 0.005 Å |
| | Angles | 0.862° |
| $B_{average}$ (no. of atoms) | Protein | 29.5 (931) |
| | Water | 35.6 (61) |
| | $Mg^{+2}$ | 26.0 (1) |

[a] $R_{merge} = \Sigma | I - < I > | / \Sigma I$
[b] PHASES output
[c] Refinement is against λ1 data.
[d] Phasing power = $[\Sigma|F_H|^2/\Sigma (F_{PH.obs} - F_{PH.calc})^2]^{1/2}$.
[e] $R_{free}$ is calculated for 10% of the data.
[f] Around 600 reflections that are close to two outermost ice rings are excluded from refinement.

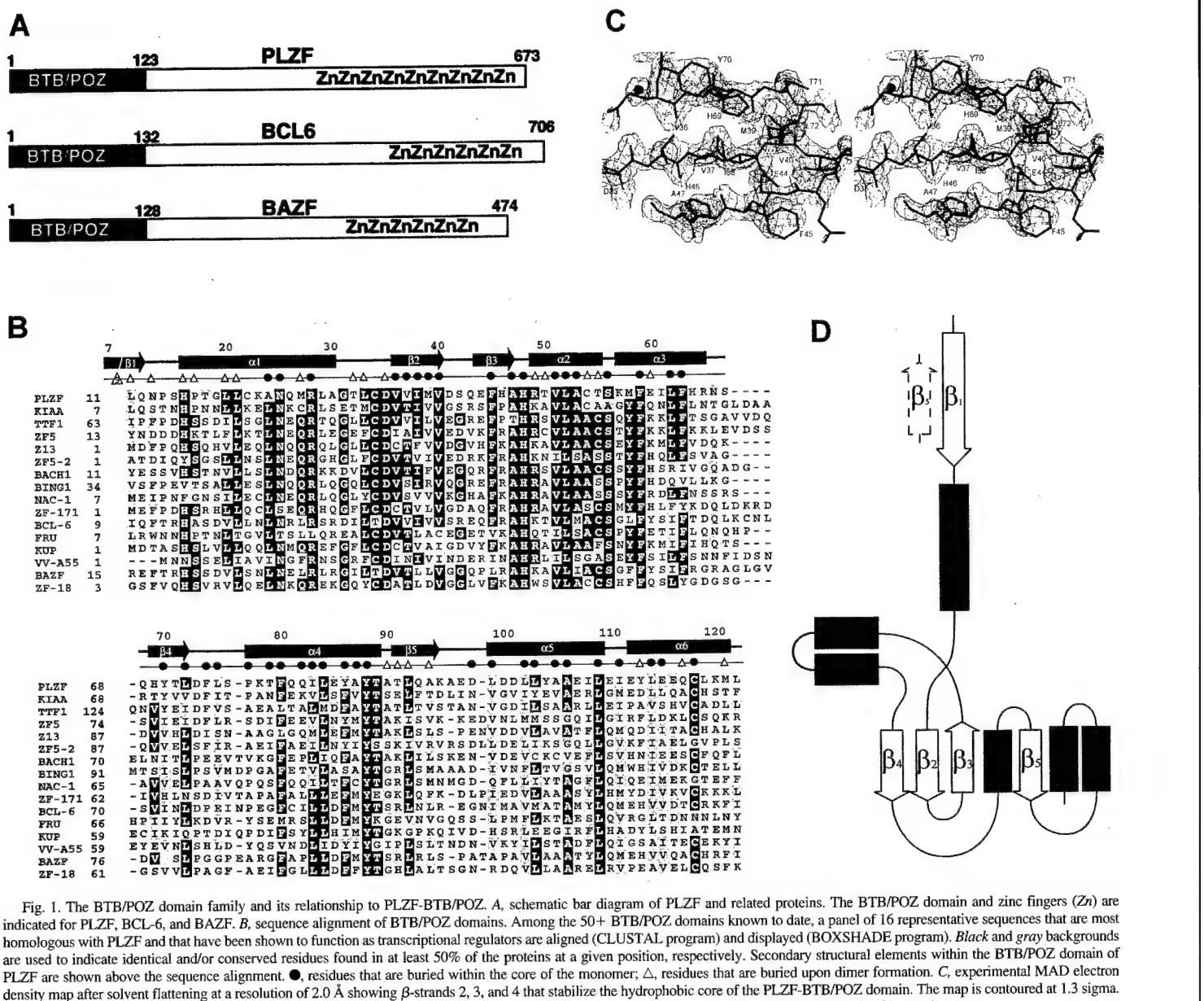



Fig. 1. The BTB/POZ domain family and its relationship to PLZF-BTB/POZ. *A*, schematic bar diagram of PLZF and related proteins. The BTB/POZ domain and zinc fingers (*Zn*) are indicated for PLZF, BCL-6, and BAZF. *B*, sequence alignment of BTB/POZ domains. Among the 50+ BTB/POZ domains known to date, a panel of 16 representative sequences that are most homologous with PLZF and that have been shown to function as transcriptional regulators are aligned (CLUSTAL program) and displayed (BOXSHADE program). *Black* and *gray* backgrounds are used to indicate identical and/or conserved residues found in at least 50% of the proteins at a given position, respectively. Secondary structural elements within the BTB/POZ domain of PLZF are shown above the sequence alignment. ●, residues that are buried within the core of the monomer; △, residues that are buried upon dimer formation. *C*, experimental MAD electron density map after solvent flattening at a resolution of 2.0 Å showing β-strands 2, 3, and 4 that stabilize the hydrophobic core of the PLZF-BTB/POZ domain. The map is contoured at 1.3 sigma. *D*, topology diagram of the BTB/POZ domain of PLZF. The β5′ strand (from the opposing subunit) is illustrated because its presence induces the β1 strand.

protein from each fraction was concentrated by deoxycholate-trichloroacetic acid precipitation (39). The precipitated protein was resuspended in 100 μl of 0.1 M NaOH. Thirty μl of the resuspended protein sample were resolved on a 12% Laemmli SDS-PAGE gel, and $PLZF_{6-123}$ proteins were visualized by fluorography.

**Transient Transfection/Luciferase Assays.** Stable expression of heterologous GAL4-$PLZF_{6-123}$ fusion proteins was confirmed by transfection in COS-1 cells. One μg of a rabbit anti-GAL4 DBD polyclonal IgG (Santa Cruz Biotech) was used to detect expression of GAL4 fusion proteins by immunoprecipitation of [$^{35}$S]methionine-labeled cell extracts (40). All transcription/luciferase assays were done in NIH/3T3 cells as described elsewhere (40).

## RESULTS AND DISCUSSION

**The Overall Structure of the PLZF-BTB/POZ Dimer Is Unique.** The BTB/POZ domain of PLZF (PLZF-BTB/POZ) crystallizes as adimer in space group I222 with one protomer per asymmetric unit with a crystallographic 2-fold passing through the functional dimer. The final model contains residues 7–122 of PLZF, which includes the complete evolutionarily conserved BTB/POZ domain (Fig. 1, *A* and *B*). Each subunit is comprised of 5 β-strands and 6 α-helices that associate to form a globular dimer (Figs. 1*D* and 2). The dimer has approximate dimensions of 20 Å × 30 Å × 60Å and has the appearance of a "butterfly," with each subunit forming most of each of the wings (Fig. 2, *A* and *B*). The β strands lie on the top and bottom of each of the subunits, and the helices fill the interior and flank the sides of the dimer. There are two pronounced clefts at the top and bottom of the dimer, where the two subunits intersect (Fig. 2*C*). The smaller upper cleft has a relatively shallow groove, whereas the larger lower cleft is 20 Å long with a width and depth of 5 and 6 Å, respectively. A search through the structural database with either the monomer or dimer shows no significant homology with other protein structures, suggesting that the PLZF-BTB/POZ domain contains a novel protein fold.

**The Core of the PLZF-BTB/POZ Domain Is Conserved within the BTB/POZ Domain Family.** Most of the secondary structural elements of the PLZF-BTB/POZ subunit contribute to stabilization of the hydrophobic core (Fig. 3*A*). The only exceptions are strands β1 and β5 and the $NH_2$ terminus of helix α1, which contribute largely to

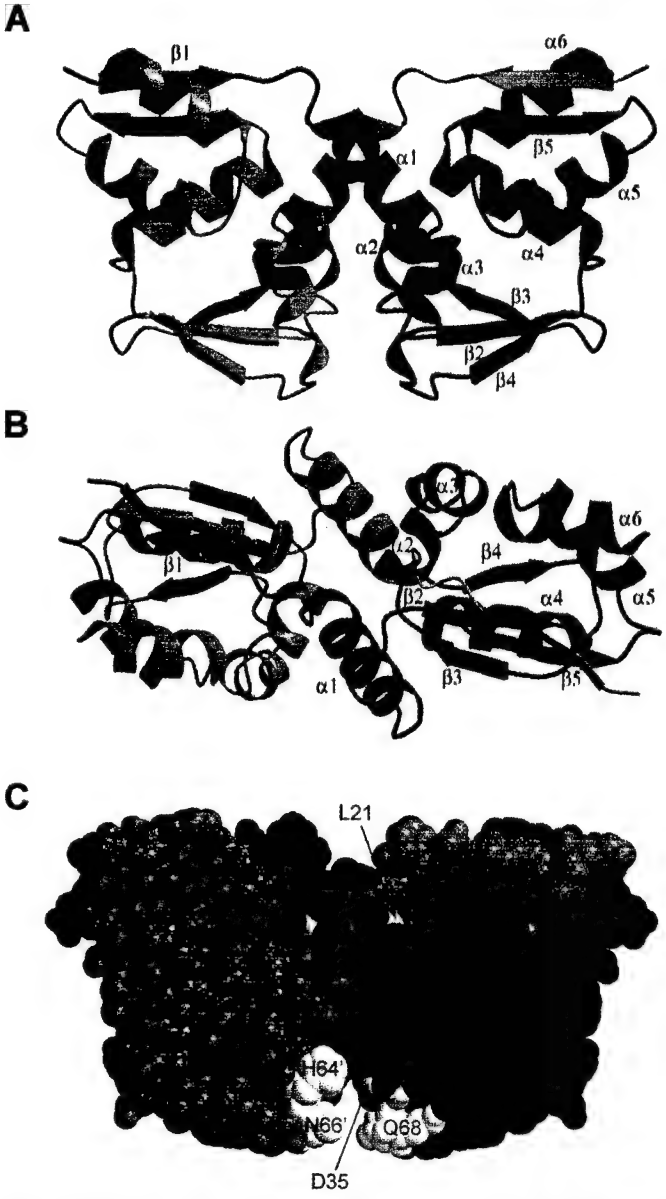



Fig. 2. Structure of the PLZF-BTB/POZ dimer. *A*, a schematic of PLZF-BTB/POZ viewed perpendicular to the 2-fold axis of the dimer. The primary subunit is shown in *green*, and the primed subunit is shown in blue. *B*, a schematic of PLZF-BTB/POZ viewed parallel to the 2-fold axis of the dimer. *C*, a space-filling model of the PLZF-BTB/POZ dimer. The position of single amino acid substitutions that were biochemically characterized in this study ($Leu^{21}$, $Asp^{35}$, $His^{64}$, $Asn^{66}$, and $Gln^{68}$) are shown in *yellow*.

dimerization (Fig. 3, *B* and *C*). $Val^{36}$, $Val^{37}$, $Ile^{38}$, $Met^{39}$, and $Val^{40}$ from $\beta 2$; $Phe^{45}$ and $Ala^{47}$ from $\beta 3$; and $Tyr^{70}$ and $Leu^{72}$ from $\beta 4$ form the floor of the hydrophobic core (Fig. 1*C*). The sides of the hydrophobic core are stabilized by a large number of residues in the helical segments. In particular, residues from helix $\alpha 2$ ($Val^{52}$, $Leu^{53}$, and $Ala^{54}$), from helix $\alpha 3$ ($Phe^{59}$, $Leu^{62}$, and $Phe^{63}$), from helix $\alpha 4$ ($Pro^{77}$, $Thr^{79}$, $Phe^{80}$, $Gln^{82}$, $Ile^{83}$, $Leu^{84}$, $Tyr^{86}$, $Ala^{87}$, and $Tyr^{88}$), and from helix $\alpha 5$ ($Leu^{99}$, $Leu^{102}$, $Leu^{103}$, $Ala^{105}$, $Ala^{106}$, and $Leu^{109}$) make substantial contributions to the core (Figs. 1*A* and 3).

The majority of residues that stabilize the core of the PLZF-BTB/POZ domain are highly conserved within the family of BTB/POZ proteins (Figs. 1*B* and 3). This conservation appears to extend beyond the subclass of BTB/POZ domain proteins that function as transcriptional regulators (7). In particular, $His^{48}$ in the loop between $\beta 3$ and $\alpha 2$, $Leu^{52}$ from $\alpha 2$, $Phe^{59}$ from $\alpha 3$, $Tyr^{88}$ from $\alpha 4$, and $Leu^{109}$ from $\alpha 5$ are conserved in virtually all BTB/POZ domains. Other residues that stabilize the core are also highly conserved among BTB/POZ proteins (Fig. 1*B*). Taken together, the conservation of residues that stabilize the core of the PLZF-BTB/POZ subunit strongly suggests that other BTB/POZ domains have a homologous subunit tertiary fold.

**Dimerization by PLZF-BTB/POZ Implicates a Conserved Mode of Dimerization by Other BTB/POZ Domain Proteins.** Previous studies in our laboratory have used a variety of biophysical techniques to show that the BTB/POZ domain of PLZF forms a dimer with an apparent $K_d < 200$ nM (8). The structure of the PLZF-BTB/POZ dimer is consistent with the high degree of dimer stability. Overall, there are 23 residues from each subunit that contribute to dimer formation, forming a solvent excluded surface of 2400 $Å^2$ for the dimer (Figs. 1*B* and 2*C*).

The principle dimer contacts between the PLZF-BTB/POZ subunits are mediated by the $\beta 1$ and $\beta 5$ strands and the $\alpha 1$ helix, which are the only secondary structural elements that do not contribute significantly to the subunit core (Fig. 3, *B* and *C*). The $\beta 1$ strand and the $NH_2$

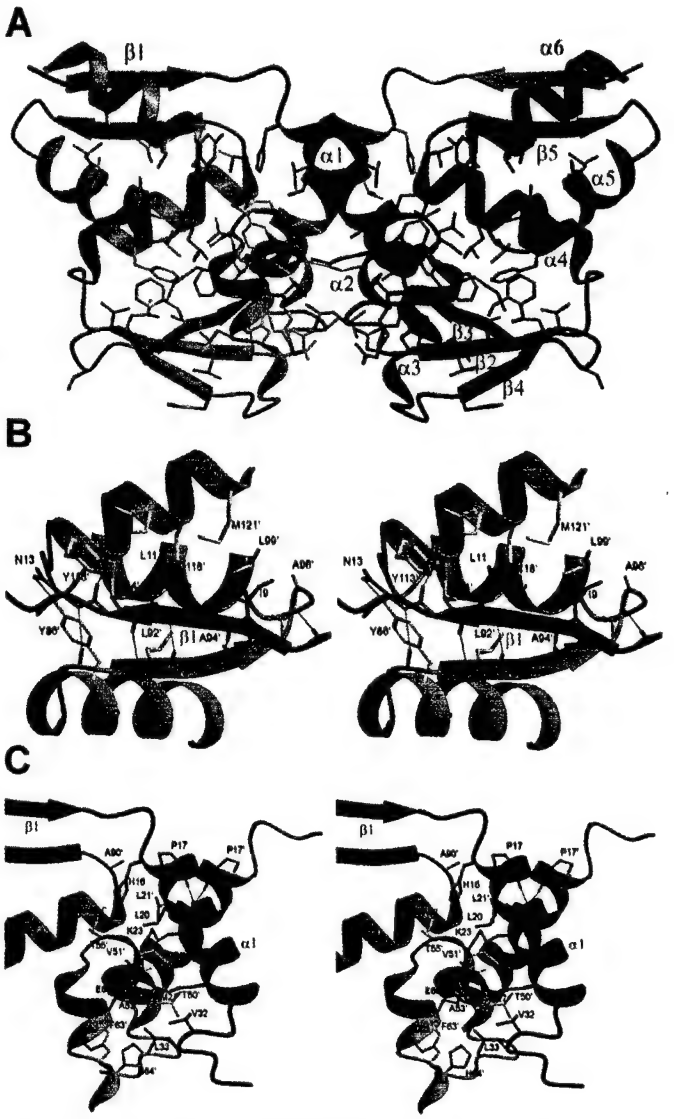



Fig. 3. Details of the BTB/POZ domain dimer. *A*, BTB/POZ conserved residues mapped onto the PLZF-BTB/POZ structure. Highly conserved residues (shaded *black* in Fig. 1*B*) are mapped onto a schematic of the PLZF-BTB/POZ dimer, illustrating that the majority of conserved residues function to stabilize the core of the monomer and the protein dimer. *B*, $\beta 1$ strand-mediated interactions that stabilize the dimer. Residues that are associated with dimer stability are indicated as *light shaded colored* side chains superimposed onto a schematic of the protein dimer. *C*, $\alpha 1$ loop-mediated interactions that stabilize the dimer.

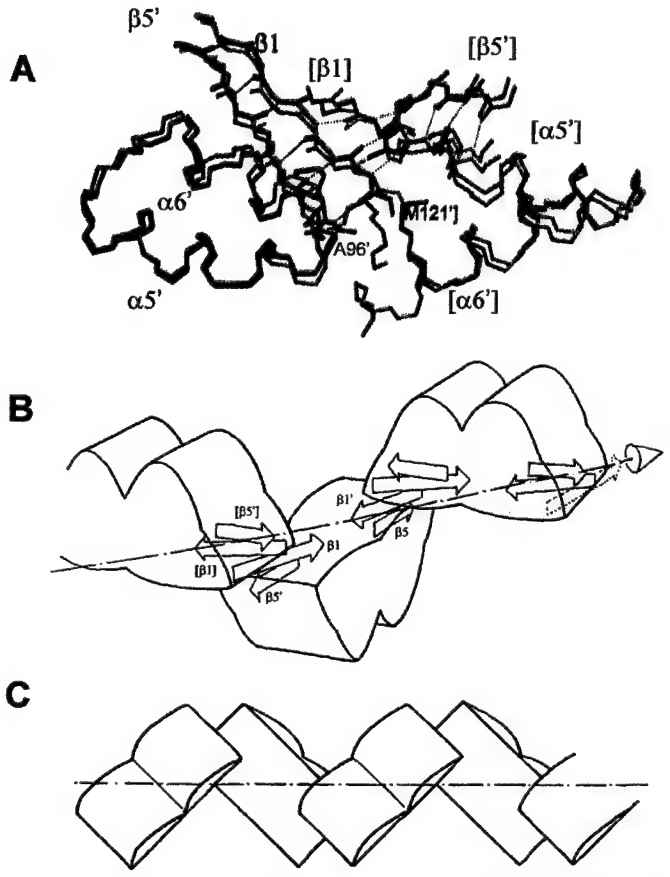



Fig. 4. Dimer-dimer interactions in the PLZF-BTB/POZ crystals. *A*, a backbone trace of the protein regions that mediate dimer-dimer interactions in the crystals are shown. The primary PLZF-BTB/POZ dimer is shown in *blue*, and the symmetry-related dimer is shown in *green* (*brackets*, secondary structural elements). Backbone hydrogen bonds within the same dimer are shown in a *light shade* of the respective color, and hydrogen bonds between dimers are shown in *red*. Ala96 and Met121, which make van der Waals interactions at the dimer interface, are also indicated as *red* side chains. A superimposition of the dimer-dimer interaction mediated in the PLZF-BTB/POZ crystals (RMS deviation for all atoms is 1.3 Å$^2$) described by Ahmad *et al.* (42) is shown in *gray*. *B* and *C*, schematic representations showing the organization of the proposed higher order BTB/POZ domain oligomers.

terminus of the $\alpha$1 helix intrude into the opposing subunit by flanking the core above ($\beta$1 and $\beta$5 strands) and in the central portion of the butterfly ($\alpha$1 helix). For simplicity of discussion, we will refer to the symmetry-related subunit of the dimer with a primed (′) designation. The $\beta$1 strand is wedged between the $\beta$5′ strand and the $\alpha$6′ helix of the opposing subunit, making sheet interactions with the $\beta$5′ strand and van der Waals and hydrogen-bond interactions with other regions of the primed subunit (Fig. 4*B*). Specifically, Ile$^9$ makes van der Waals interactions with Leu$^{99}$ of $\alpha$5′, Ala$^{96'}$ in the loop between $\beta$5′ and $\alpha$6′ and Met$^{121}$ of $\alpha$6′; Ile$^{11}$ makes van der Waals interactions with Tyr$^{113}$, Leu$^{114}$, Glu$^{117}$, and Cys$^{118}$ of helix $\alpha$6′; and the N$\delta$2 of Asn$^{13}$ hydrogen bonds to the side-chain hydroxyl of Tyr$^{86}$ in helix $\alpha$4′.

The $\alpha$1 helix and the proceeding loop is situated between the $\alpha$1′ and $\alpha$2′ helices and makes mostly van der Waals interactions with residues within these helices as well as with the $\alpha$3′ and $\alpha$4′ helices (Fig. 3*C*). Specifically, $\alpha$1-helix residues His$^{16}$, Pro$^{17}$, Leu$^{20}$, Leu$^{21}$, and Met$^{27}$ make a series of van der Waals interactions with residues from helices $\alpha$1′, $\alpha$2′, $\alpha$3′, and $\alpha$4′. In particular, Leu$^{21}$ from $\alpha$1′, Cys54 from $\alpha$2′, and Ala$^{90}$ from the loop between $\alpha$4′ and $\beta$5′ play important roles in stabilizing the dimer. Thr$^{32}$, Leu$^{33}$, and Asp$^{35}$ in the loop proceeding helix $\alpha$1 also contribute to dimer stability. His$^{64}$ from $\alpha$3′ plays a particularly important role in this regard.

Nearly 80% of the residues that stabilize the PLZF-BTB/POZ dimer show conservation within the BTB/POZ domain family (Fig. 1*A*). In particular, His$^{16}$ and Leu$^{21}$ in helix $\alpha$1 and Asp$^{35}$ in the loop proceeding the $\alpha$1 helix are highly conserved and play important roles in dimer stability. All but the aspartic acid play a hydrophobic role in dimer stability. Asp$^{35}$, in contrast, makes a direct hydrogen bond to Arg$^{49}$ (moderately conserved within the BTB/POZ family) and a water-mediated hydrogen bond to Asp$^{35'}$. Taken together, the degree of conservation within residues that stabilize dimer formation suggests that BTB/POZ domains from otherwise unrelated proteins will form dimers with similar quaternary arrangements. Correlating well with our findings, there are several BTB/POZ domains that have been shown to form homodimers. Among them are PLZF (8), ZID (6), Ttk (6), bab (9), and BAZF (11). Indeed, the BTB/POZ domain appears to be an ideally suited dimerization module.

**Dimer-Dimer Interactions in the Crystals Implicates a Propensity for Higher Order Oligomerization by BTB/POZ Domains *in Vivo*.** Comparison of the PLZF-BTB/POZ structure derived here with that of the recently published PLZF-BTB/POZ structure determined by Ahmad *et al.* (42) shows a high degree of structural similarity between the protein dimers with an RMS deviation between all atoms of 1.1 Å$^2$. Strikingly, this comparison also shows that although the two structures were obtained in different crystal lattice environments, both show structurally homologous dimer-dimer interactions in the crystal lattice (Fig. 4*A*). These dimer-dimer contacts bury a total of 1200 Å$^2$ of solvent excluded surface and is largely mediated by 2 4-stranded antiparallel $\beta$-sheet involving $\beta$1 and $\beta$5′ from one dimer with the corresponding segments of the symmetry related dimer. In addition, Ala$^{96}$ in the loop between $\beta$5 and $\alpha$5 makes a van der Waals contact with Met$^{121}$ at the COOH-terminus of $\alpha$6 in the symmetry related dimer. The $\alpha$6-mediated interactions at the dimer-dimer interface are somewhat more extensive in the PLZF-BTB/POZ structure described by Ahmad *et al.* (42) because their $\alpha$6 helix contains an additional turn of secondary structure. It is interesting to note that the mode of dimer-dimer interaction observed in the crystals does not prohibit the formation of extended dimer-dimer interactions, which would result in the formation of even higher order multimers (Fig. 4, *B* and *C*). Taken together, these observations suggest that the BTB/POZ domain may mediate the formation of higher-order multimers *in vivo*.

Correlating well with our findings, two recent studies have shown that the BTB/POZ domain of the GAGA transcription factor directly mediates the formation of higher-order oligomers to bind multiple GAGA sites that are found in natural target promoters *in vivo* (41, 43). Moreover, the formation of these higher-order oligomers have been shown to be correlated with the cooperative nature of GAGA transcription factor binding to multiple DNA-binding sites and correlated with the finding that this cooperativity is strictly dependent on the presence of the GAGA BTB/POZ domain. Interestingly, natural GAGA promoters display a large degree of variability between GAGA sites, also correlating well with the relatively flexible dimer-dimer interactions seen in the crystals.

Our findings of higher-order BTB/POZ oligomers may also explain other studies that find that the BTB/POZ domains from some proteins mediate specific hetero-oligomers. For example, Ttk can form oligomers with itself and with the GAGA protein but not with the BTB/POZ region of ZID (6). In addition, BAZF has been shown to form oligomers with itself and with BCL6 (11). Taken together, the BTB/POZ domain appears to be an ideally suited module for both homo and hetero protein multerimerization.

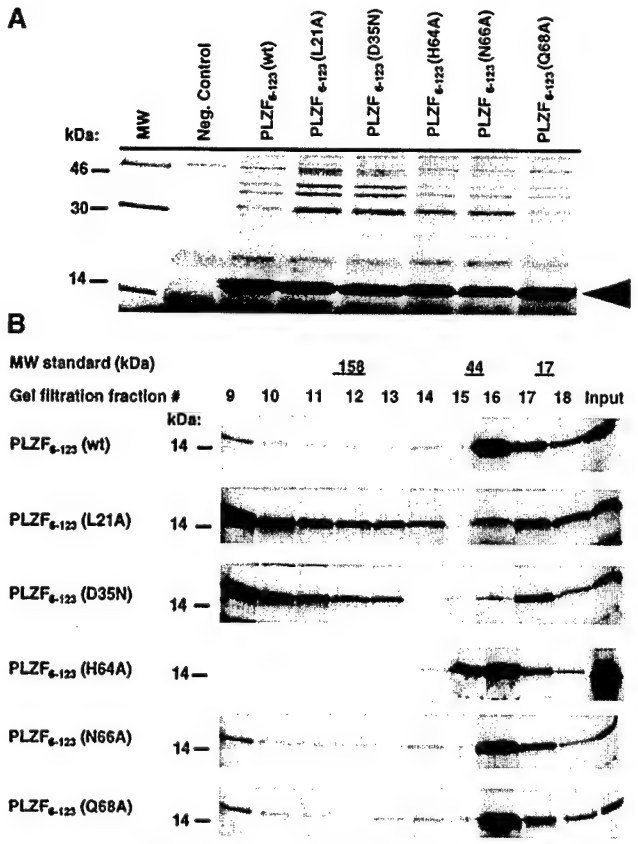



Fig. 5. Effects of single amino acid substitutions on the dimerization properties of the PLZF BTB/POZ domain. *A, in vitro* translation of wild-type and $PLZF_{6-123}$-BTB/POZ proteins containing single amino acid substitutions (L21A, D35N, H64A, N66A, and Q68A). Molecular weight standards and a negative control in which a translation reaction was done in the absence of RNA polymerase are indicated. *B,* Superdex 200 HR 10/30 gel filtration fractions of *in vitro*-translated, $^{35}S$-labeled wild-type and mutant $PLZF_{6-123}$-BTB/POZ proteins were resolved in SDS-PAGE gels and visualized by fluorography. Elution of molecular weight standards was determined in a parallel gel filtration experiment under exactly the same conditions. Fractions with peak levels of molecular weight standards are indicated at the *top* of the figure.

**Biochemical Analysis of PLZF Proteins Harboring Site-directed Mutations Shows That Dimerization Is Required for Transcriptional Repression.** The recent study by Ahmad *et al.* (42) has suggested that a pronounced cleft along the bottom of the PLZF dimer (Fig. 2*C*) may be a site of corepressor interaction. To directly test this hypothesis and to directly probe the functional significance of dimer formation by PLZF, we carried out site-directed mutagenesis coupled with biochemical analysis of these mutant proteins. We prepared five different site-directed mutations that fell into two subgroups (Fig. 2*C*). The first two (L21A and D35N) involved residues that play critical roles in dimerization; and the second group (H64A, N66A, and Q68A) involved residues in the pronounced cleft at the base of the PLZF-BTB/POZ domain dimer. Each of the mutants were tested for both dimerization and transcriptional repression properties. To evaluate the effects of these mutations on the dimerization of the BTB/POZ domain, the wild-type $PLZF_{6-123}$ and each mutant described herein were *in vitro* translated and then subjected to gel filtration (Fig. 5). Consistent with expectations, the wild-type $PLZF_{6-123}$ protein, as well as each of the proteins harboring mutations in the cleft region, eluted from the sizing column at a molecular weight consistent with a dimeric PLZF-BTB/POZ species. In addition, proteins harboring mutations in the dimerization interface of PLZF-BTB/POZ eluted in two peaks, one near the void volume ($M_r \leq 670{,}000$), indicative of protein aggregates and another at a position consistent with a monomeric PLZF-BTB/POZ domain ($M_r$ 17,000). These results indicated that the L21A and D35N mutant proteins were defective in dimer formation.

To evaluate any effects these mutations may have on the transcriptional repression function of the PLZF BTB/POZ domain, heterologous fusions between the DNA binding domain of GAL4 (amino acids 1–147) and each of the BTB/POZ domain mutants described herein were created (Fig. 6). Each fusion protein demonstrated stable and comparable nuclear-localized expression in COS-1 cells (Fig. 6*B*), suggesting comparable DNA-binding properties for each of our protein fusions. The wild-type GAL4-$PLZF_{6-123}$ BTB/POZ domain and each mutant version thereof were cotransfected with 2 TK-luciferase reporter downstream of four copies of the GAL4 consensus UAS in NIH/3T3 cells (Fig. 6*A*). The wild-type GAL4-$PLZF_{6-123}$ protein demonstrated significant repression of the luciferase reporter gene activity (Fig. 6*C*). Expression of the of the D35N mutant protein, which fails to dimerize properly (Fig. 5*B*), demonstrated a significantly impaired repression function (Fig. 6*C*). To a lesser extent, the repression function of the L21A mutant protein was also impaired (Fig. 6*C*). All other mutants (H64A, N66A, and Q68A) demonstrated negligible effects on the repression function of the BTB/POZ domain of PLZF (Fig. 6*C*). Significantly, these data suggest that the repression function of the PLZF-BTB/POZ domain is strictly dependent upon

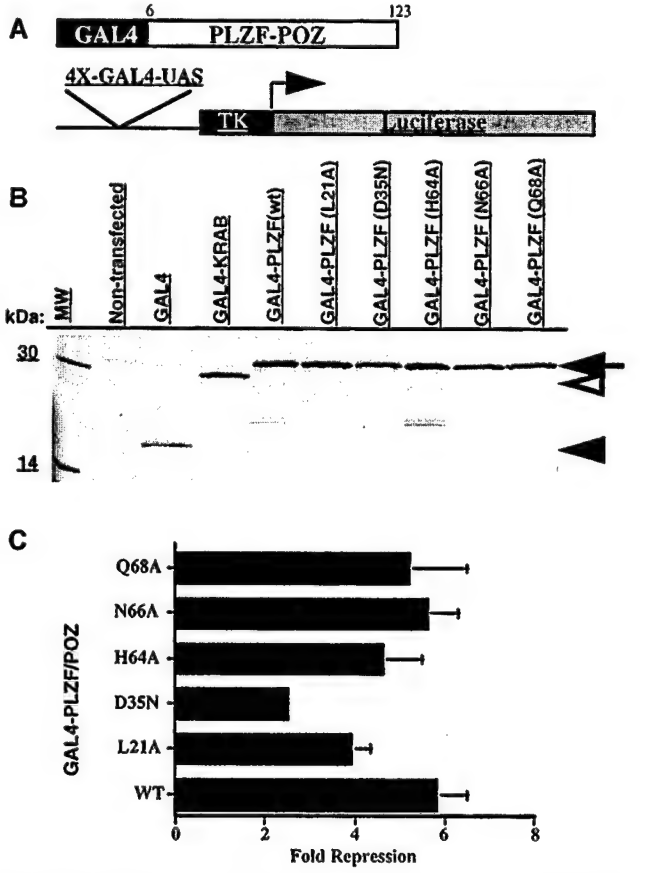



Fig. 6. Effects of single amino acid substitutions on the transcriptional repression properties of the PLZF BTB/POZ domain. *A,* schematic diagram illustrating the plasmids used in transfection of NIH/3T3 cells. GAL4 fusion proteins of the wild-type and $PLZF_{6-123}$ proteins containing single amino acid substitutions are shown. *B,* similar expression levels of GAL4-$PLZF_{6-123}$ fusion proteins was confirmed by transient transfection in COS-1 cells, followed by immunoprecipitation of [$^{35}S$]methionine-labeled cell extracts with anti-GAL4 DBD polyclonal IgG. *Closed arrowhead,* GAL4 protein; *open arrowhead,* GAL4-KRAB protein as a positive control. *Arrow,* wild-type and mutant GAL4- $PLZF_{6-123}$ fusion proteins. Molecular weight standards are indicated. *C,* the relative levels of transcriptional repression of the GAL4-$PLZF_{6-123}$ fusion proteins are indicated in terms of fold-repression. Values were calculated by comparing normalized luciferase activities with cells transfected in the absence of a GAL4 effector protein. Cotransfection with a pcDNA3-$\beta$-galactosidase expression plasmid was used to normalize all luciferase values for transfection efficiency. *Bars,* SEs observed for three independent transfections (each performed in duplicate).

homodimerization. Moreover, these data suggest that the cleft region of the PLZF-BTB/POZ domain may not play a significant role in the transcriptonal repression property of PLZF.

Recently, Huynh and Bardwell (10) have carried out a site-directed mutagenesis study of the BTB/POZ domain of BCL6 with the goal of disrupting corepressor function. In this study, the authors targeted $NH_2$-terminal residues that are highly conserved in proteins that showed interaction with the corepressor proteins N-CoR and SMRT and that mediated transcriptional repression (BCL6, PLZF, ZID, GAGA, and vaccinia virus SalF17R). Mutations were made to the corresponding residues of the BTB/POZ harboring protein, Protein A from the *mod(mdg4)* gene, which failed to show binding to N-CoR and SMRT and which failed to mediate transcriptional repression. A correlation of the phenotype of these mutations with the PLZF-BTB/POZ structure shows that the two mutations that disrupt corepressor binding and function (Leu to Ser at position 21, and Met-Arg to Ser-Leu at positions 27 and 28 using PLZF numbering) would also disrupt dimerization of the PLZF-BTB/POZ region (Figs. 2*C* and 3*C*). An asparagine to histidine change at position 25 (using PLZF numbering) did not have a detectable effect on N-CoR and SMRT interaction or transcriptional repression. The structure of the PLZF-BTB/POZ domain shows that $Asp^{25}$ is packed on the interior of the protein, thus playing an important role in protein folding (Fig. 3*A*). Modeling of a histidine in position 25 shows that it could carry out a comparable role in BTB/POZ domain stability. Taken together, our results and the results of Huynh and Bardwell (10) underscore the importance of BTB/POZ domain dimerization for N-CoR and SMRT-mediated transcriptional repression.

The precise mode of corepressor recognition by PLZF must await the structure of an appropriate PLZF-BTB/POZ/corepressor complex. Nonetheless, the structural and functional information provided here provides a framework from which to use structure/function analysis to better understand the N-CoR and SMRT binding and repression properties of the BTB/POZ domain of PLZF. Moreover, the information provided here provides a conceptual and structural scaffold from which to design PLZF-specific inhibitory molecules that may target the dimerization and/or interaction of PLZF with corepressor proteins to be used to treat APL patients harboring the PLZF-RARα translocation.

## ACKNOWLEDGMENTS

We thank Craig Ogata and his staff for access to and help on beamline X4A at NSLS; David Speicher (Wistar Institute microchemistry core facility) for mass spectrometry and $NH_2$-terminal sequencing analysis of SeMet-derivatized PLZF-BTB/POZ protein; and Greg Van Duyne, Mitch Lewis, Dan King, Ravi Venkataramani, Yi Mo, Stacey Benson, and Kunchithapadam Swaminathan for useful discussions. Coordinates for the PLZF-BTB/POZ structure can be obtained from the Research Collaboratory for Structural Bioinformatics (RCSB009519).

## REFERENCES

1. Chen, Z., Brand, N. J., Chen, A., Chen, S-J., Tong, J-H., Wang, Z. Y., Waxman, S., and Zelent, A. Fusion between a novel *Kruppel*-like zinc finger gene and the retinoic acid receptor-α locus due to a variant t(11;17) translocation associated with acute promyelocytic leukemia. EMBO J., *12:* 1161–1167, 1993.
2. Rowley, J. D., Golomb, H. M., and Dougherty, C. 15/17 translocation: a consistent chromosomal change in acute promyelocytic leukaemia [letter]. Lancet, *1:* 549–550, 1977.
3. Zelent, A. Translocation of the *RARα* locus to the PML of *PLZF* gene in acute promyelocytic leukemia. Br. J. Haematol., *86:* 451–460, 1994.
4. Licht, J. D., Chomienne, C., Goy, A., Chen, A., Scott, A. A., Head, D. R., Michaux, J. L., Wu, Y., DeBlasio, A., Miller, W. H., Jr., Zelenetz, A. D., Willman, C. L., Chen, Z., Chen, S-J., Zelent, A., Macintyre, E., Veil, A., Cortes, J., Kantarjian, H., and Waxman, S. Clinical and molecular characterization of a rare syndrome of acute promyelocytic leukemia associated with translocation (11;17). Blood, *85:* 1083–1094, 1995.
5. Zollman, S., Godt, D., Prive, G. G., Couderc, J-L., and Laski, F. The BTB domain, found primarily in zinc finger proteins, defines an evolutionarily conserved family that includes several developmentally regulated genes in *Drosophila*. Proc. Natl. Acad. Sci. USA, *91:* 10717–10721, 1994.
6. Bardwell, V. J., and Treisman, R. The POZ domain: a conserved protein-protein interaction motif. Genes Dev., *8:* 1664–1677, 1994.
7. Albagli, O., Dhordain, P., Deweindt, C., Lecocq, G., and Leprince, D. The BTB/POZ domain: a new protein-protein interaction motif common to DNA- and actin-binding proteins. Cell Growth Differ., *6:* 1193–1198, 1995.
8. Li, X., Lopez-Guisa, J. M., Ninan, N., Weiner, E. J., Rauscher, F. J., III, and Marmorstein, R. Overexpression, purification, characterization, and crystallization of the BTB/POZ domain from the PLZF oncoprotein. J. Biol. Chem., *272:* 27324–27329, 1997.
9. Chen, W., Zollman, S., Couderc, J. L., and Laski, F. A. The BTB domain of *bric a brac* mediates dimerization *in vitro*. Mol. Cell. Biol., *15:* 3424–3429, 1995.
10. Huynh, K. D., and Bardwell, V. J. The BCL-6 POZ domain and other POZ domains interact with the co-repressors N-CoR and SMART. Oncogene, *17:* 2473–2484, 1998.
11. Okabe, S., Fukuda, T., Ishibashi, K., Kojima, S., Okada, S., Hatano, M., Ebara, M., Saisho, H., and Tokuhisa, T. BAZF, a novel BCL6 homolog, functions as a transcriptional repressor. Mol. Cell. Biol., *18:* 4235–4244, 1998.
12. Numoto, M., Niwa, O., Kaplan, J., Wong, K. K., Merrell, K., Kamiya, K., Yanagihara, K., and Calame, K. Transcriptional repressor ZF5 identifies a new conserved domain in zinc finger proteins. Nucleic Acids Res., *21:* 3767–3775, 1993.
13. Albagli, O., Dhordain, P., Bernardin, F., Quief, S., Kerckaert, J-P., and Leprince, D. Multiple domains participate in distance-independent LAZ3/BCL6-mediated transcriptional repression. Biochem. Biophys. Res. Commun., *220:* 911–915, 1996.
14. Chang, C-C., Ye, B. H., Chaganti, R. S. K., and Dalla-Favera, R. BCL-6, a POZ/zinc-finger protein, is a sequence-specific transcriptional repressor. Proc. Natl. Acad. Sci. USA, *93:* 6947–6952, 1996.
15. Seyfert, V., Allman, D., He, Y., and Staudt, L. M. Transcriptional repression by the proto-oncogene *BCL-6*. Oncogene, *12:* 2331–2342, 1996.
16. Deweindt, C., Albagli, O., Bernardin, F., Dhordain, P., Quief, S., Lantoine, D., Kerckaert, J-P., and Leprince, D. The *LAZ3/BCL6* oncogene encodes a sequence-specific transcriptional inhibitor: a novel function for the BTB/POZ domain as an autonomous repressing domain. Cell Growth Differ., *6:* 1495–1503, 1995.
17. Dhordain, P., Albagli, O., Lin, R. J., Ansieau, S., Quief, S., Leutz, A., Kerckaert, J-P., Evans, R. M., and Leprince, D. L. Corepressor SMART binds the BTB/POZ repressing domain of the LAZ3/BCL6 oncoprotein. Proc. Natl. Acad. Sci. USA, *94:* 10762–10767, 1997.
18. Hong, S-H., David, G., Wong, C-W., Dejean, A., and Privalsky, M. L. SMART corepressor interacts with PLZF and with the PML-retinoic acid receptor α (RARα) and PLZF-RARα oncoproteins associated with acute promyelocytic leukemia. Proc. Natl. Acad. Sci. USA, *94:* 9028–9033, 1997.
19. David, G., Alland, L., Hong, S-H., Wong, C-W., Depinho, R., and Dejean, A. Histone deacetylase associated with mSin3A mediates repression by the acute promyelocytic leukemia-associated PLZF protein. Oncogene, *16:* 2549–2556, 1998.
20. Lin, R. J., Nagy, L., Inoue, S., Shao, W. L., Miller, W. H., and Evans, R. M. Role of the histone deacetylase complex in acute promyelocytic leukaemia. Nature (Lond.), *391:* 811–814, 1998.
21. Grignani, F., De Matteis, S., Nervi, C., Tomassoni, L., Gelmetti, V., Cioce, M., Fanelli, M., Ruthardt, M., Ferrara, F. F., Zamir, I., Seiser, C., Lazar, M. A., Minucci, S., and Pelicci, P. G. Fusion proteins of the retinoic acid receptor-α recruit histone deacetylase in promyelocytic leukaemia. Nature (Lond.), *391:* 815–818, 1998.
22. Chen, J. D., and Evans, R. M. A transcriptional co-repressor that interacts with nuclear hormone receptors. Nature (Lond.), *377:* 454–457, 1995.
23. Heinzel, T., Lavinsky, R. M., Mullen, T-M., Soderstrom, M., Laherty, C. D., Torchia, J., Yang, W-M., Brard, G., Ngo, S. D., Davie, J. R., Seto, E., Eisenman, R. N., Rose, D. W., Glass, C. K., and Rosenfeld, M. G. A complex containing N-CoR, mSin3 and histone deacetylase mediates transcriptional repression. Nature (Lond.), *387:* 43–48, 1997.
24. Alland, L., Muhle, R., Hou, H., Potes, J., Chin, L., and Schreiber-Agus, N. Role for N-CoR and histone deacetylase in Sin3-mediated transcriptional repression. Nature (Lond.), *387:* 49–55, 1997.
25. He, L. Z., Guidez, F., Tribioli, C., Peruzzi, D., Ruthardt, M., Zelent, A., and Pandolfi, P. P. Distinct interactions of PML-RARα and PLZF-RARα with co-repressors determine differential responses to RA in APL. Nat. Genet., *18:* 126–135, 1998.
26. Guidez, F., Huag, W., Tong, J-H., Dubois, C., Balitrand, N., Waxman, S., Michaux, J. L., Martial, P., Degos, L., Chen, Z., and Chomienne, C. Poor response to all-*trans* retinoic acid therapy in a t(11;17) PLZF/RARα patient. Leukemia (Baltimore), *8:* 312–317, 1994.
27. Neidhardt, F. C., Bloch, P. L., and Smith, D. F. Culture medium for enterobacteria. J. Bacteriol., *119:* 736–747, 1974.
28. Otwinowski, Z. Oscillation data reduction program. *In:* L. Sawyer, N. Isaacs, and S. Bailey (eds.), Proceedings of the CCP4 Study Weekend: Data Collection and Processing, pp. 56–62. Warrington, United Kingdom: SERC Daresbury Laboratory, 1993.
29. Furey, W., and Swaminathan, S. PHASES-95: a program package for the processing and analysis of diffraction data from macromolecules. *In:* C. W. Carter and R. M. Sweet (eds.), Methods in Enzymology, Vol. 277, pp. 590–620. Orlando, FL: Academic Press, 1997.
30. Jones, T. A., Zou, J. Y., and Cowen, S. W. Improved methods for building protein models in electron density maps and the location of errors in these models. Acta Crystallogr. A, *47:* 110–119, 1991.

31. Brunger, A. T. X-PLOR 3.1, a system for X-ray crystallography and NMR. New Haven, CT: Yale University Press, 1992.
32. Brunger, A. T. The free R value: a novel statistical quantity assessing the accuracy of crystal structures. Nature (Lond.), *335:* 472–474, 1992.
33. Brunger, A. T., Kuriyan, J., and Karplus, M. Crystallographic R factor refinement by molecular dynamics. Science (Washington DC), *235:* 458–460, 1987.
34. Brunger, A. T., and Krukowski, A. Slow-cooling protocols for crystallographic refinement by simulated annealing. Acta Crystallogr. A, *46:* 585–593, 1990.
35. Rice, L. M., and Brunger, A. T. Torsion angle dynamics: reduced variable conformational sampling enhances crystallographic structure refinement. Proteins, *19:* 277–290, 1994.
36. Jiang, J. S., and Brunger, A. T. Protein hydration observed by X-ray diffraction: solvation properties of penicillopepsin and neuraminidase crystal structures. J. Mol. Biol., *243:* 100–115, 1994.
37. Hodel, A., Kim, S-H., and Brunger, A. T. Model bias in macromolecular crystal structures. Acta Crystallogr. A, *48:* 851–858, 1992.
38. Kleywegt, G. J., and Jones, T. A. Phi/Psi-chology: Ramachandran revisted. Curr. Biol., *4:* 1395–1400, 1996b.
39. Peterson, G. L. A simplification of the protein assay method of Lowry *et al.* which is more generally applicable. Anal. Biochem., *83:* 346–356, 1977.
40. Ryan, R. F., Schultz, D. C., Ayyanathan, K., Singh, P. B., Friedman, J. R., Fredericks, W. J., and Rauscher, F. J. KAP-1 corepressor protein interacts and colocalizes with heterochromatic and euchromatic HP1 proteins: a potential role for Kruppel-associated Box-zinc finger proteins in heterochromatin-mediated gene silencing. Mol. Cell. Biol., *19:* 4366–4378, 1999.
41. Espinas, M. L., Jimenez-Garcia, E., Vaquero, A., Canudas, S., Bernues, J., and Azorin, F. The N-terminal POZ domain of GAGA mediates the formation of oligomers that bind DNA with high affinity and specificity. J. Biol. Chem., *274:* 16461–16469, 1999.
42. Ahmad, K. F., Engel, C. K., and Prive, G. G. Crystal structure of the BTB domain from PLZF. Proc. Natl. Acad. Sci. USA, *95:* 12123–12128, 1998.
43. Katsani, K. R., Hajibagheri, M. A. N., and Verrijzer, C. P. Co-operative DNA binding by GAGA transcription factor requires the conserved BTB/POZ domain and reorganizes promoter topology. EMBO J., *18:* 698–708, 1999.

MOLECULAR AND CELLULAR BIOLOGY, June 1999, p. 4366–4378 Vol. 19, No. 6
0270-7306/99/$04.00+0


# KAP-1 Corepressor Protein Interacts and Colocalizes with Heterochromatic and Euchromatic HP1 Proteins: a Potential Role for Krüppel-Associated Box–Zinc Finger Proteins in Heterochromatin-Mediated Gene Silencing

ROBERT F. RYAN,[1] DAVID C. SCHULTZ,[1] KASIRAJAN AYYANATHAN,[1] PRIM B. SINGH,[2] JOSH R. FRIEDMAN,[1] WILLIAM J. FREDERICKS,[3] AND FRANK J. RAUSCHER III[1]*

*The Wistar Institute, Philadelphia, Pennsylvania[1]; Department of Development and Reproduction, The Roslin Institute, Edinburgh, United Kingdom[2]; and Onyx Pharmaceuticals, Richmond, California[3]*

Received 28 December 1998/Returned for modification 8 February 1999/Accepted 29 February 1999

**Krüppel-associated box (KRAB) domains are present in approximately one-third of all human zinc finger proteins (ZFPs) and are potent transcriptional repression modules. We have previously cloned a corepressor for the KRAB domain, KAP-1, which is required for KRAB-mediated repression in vivo. To characterize the repression mechanism utilized by KAP-1, we have analyzed the ability of KAP-1 to interact with murine (M31 and M32) and human (HP1α and HP1γ) homologues of the HP1 protein family, a class of nonhistone heterochromatin-associated proteins with a well-established epigenetic gene silencing function in *Drosophila*. In vitro studies confirmed that KAP-1 is capable of directly interacting with M31 and hHP1α, which are normally found in centromeric heterochromatin, as well as M32 and hHP1γ, both of which are found in euchromatin. Mapping of the region in KAP-1 required for HP1 interaction showed that amino acid substitutions which abolish HP1 binding in vitro reduce KAP-1 mediated repression in vivo. We observed colocalization of KAP-1 with M31 and M32 in interphase nuclei, lending support to the biochemical evidence that M31 and M32 directly interact with KAP-1. The colocalization of KAP-1 with M31 is sometimes found in subnuclear territories of potential pericentromeric heterochromatin, whereas colocalization of KAP-1 and M32 occurs in punctate euchromatic domains throughout the nucleus. This work suggests a mechanism for the recruitment of HP1-like gene products by the KRAB-ZFP–KAP-1 complex to specific loci within the genome through formation of heterochromatin-like complexes that silence gene activity. We speculate that gene-specific repression may be a consequence of the formation of such complexes, ultimately leading to silenced genes in newly formed heterochromatic chromosomal environments.**

Regulation of gene expression at the level of transcription initiation by sequence-specific DNA binding proteins has emerged as one of the most important modes of metazoan development and homeostasis (6). The population of transcription factors that are active in the cell nucleus largely dictates the transcriptional output from the nucleus and hence the proliferative or differentiated phenotype of the cell. The dominant theme that has emerged from the study of eukaryotic transcriptional regulatory proteins is that they are highly modular in architecture, with independent, functionally separable domains mediating nuclear localization, sequence-specific DNA binding, hetero- or homo-oligomerization, activation, and repression of transcription. Recently, much effort has been expended to understand how activation and repression domains transmit the signal for modulation of transcription from a DNA-bound protein to the RNA synthesis machinery.

Studies aimed at understanding the mechanisms of transcription repression have been greatly aided by the realization that the domains which mediate repression are often highly conserved amino acid sequence motifs which occur in one or more families of proteins with common DNA binding domains. Examples of these domains include BTB/POZ, WRPW, SNAG, and Krüppel-associated box (KRAB) (2, 5, 11, 20). We have focused on the KRAB domain as a model system for analysis of conserved repression modules (17, 34). The KRAB repression domain was originally identified in humans as a conserved amino acid sequence motif at the amino termini of proteins which contain multiple TFIIIA/Krüppel class $Cys_2$-$His_2$ ($C_2H_2$) zinc fingers in their COOH termini (5) and has been identified in frog, rodent, and human zinc finger proteins (ZFPs) (for a review, see reference 24). It has been estimated that between 300 and 700 human genes encode $C_2H_2$ zinc finger proteins (28), one-third of which are predicted to contain KRAB domains (5); accordingly, these genes have been designated the KRAB-ZFP family. The KRAB domain homology consists of approximately 75 amino acids (aa) which can function as a potent transferable DNA binding-dependent repression module. Moreover, more than 10 independently encoded KRAB domains have been demonstrated to be potent repressors, and substitutions of conserved residues within this domain abolish repression activity (34). These observations suggested that transcription repression is a common property of this domain (for a review, see reference 24).

Repression mediated by DNA binding proteins has been shown to proceed via several mechanisms, including histone deacetylation (21, 51, 55), template heterochromatinization (19, 38, 45), and direct interaction with components of the transcription machinery (3, 4, 14, 15, 43, 53). Like activators, many eukaryotic repressor proteins recruit specific corepressors via protein-protein interactions, and these interactions appear to be necessary for template silencing. To identify the mechanism(s) of KRAB-ZFP-mediated repression, we previ-

* Corresponding author. Mailing address: The Wistar Institute, 3601 Spruce St., Philadelphia, PA 19104. Phone: (215) 898-0995. Fax: (215) 898-3929. E-mail: Rauscher@wista.wistar.upenn.edu.

ously identified and cloned the gene encoding a KRAB domain binding protein, KAP-1, which shows all the hallmarks of being a universal corepressor for the KRAB domain (17). KAP-1 was subsequently identified by other investigators by using yeast two-hybrid screens as transcription intermediary factor 1β (TIF1β) and KRIP-1 (27, 30, 36). KAP-1 is a 97-kDa nuclear phosphoprotein whose primary amino acid sequence displays a number of interesting structural motifs. The RING finger, B boxes (β1 and β2), and a coiled-coil region at the amino terminus collectively constitute the KRAB interaction, or RBCC, domain (17, 36, 42). Carboxy terminal to this constellation of motifs appears a relatively novel stretch of amino acids, a plant homeodomain (PHD) finger, and a bromodomain, which likely represent at least two or more independent repression domains (17, 42).

A number of lines of evidence have suggested that KAP-1 plays a key role in mediating KRAB domain repression: (i) KAP-1 binds to multiple KRAB repression domains both in vitro and in vivo, (ii) KRAB domain mutations which abolish repression decrease or eliminate the interaction with KAP-1, (iii) overexpression of KAP-1 enhances KRAB-mediated repression in a manner dependent on the presence of the RBCC domain, and (iv) heterologous fusions between KAP-1 and a DNA binding domain can potentiate repression (17, 36, 42). Finally, the KRAB domain does not exhibit repression activity in cells which lack KAP-1 protein (42). These results support a model in which KRAB-ZFPs bind a gene in a DNA sequence-specific manner and repress transcription of the bound gene by recruiting the KAP-1 corepressor. The next question is: what are the molecules downstream of the KAP-1 corepressor which mediate transcription repression?

Clues to the nature of the downstream players in the repression pathway have come from the analysis of KAP-1 homologues and orthologues. In a functional screen for nuclear hormone receptor coactivators, TIF1α was cloned and shown to be similar in overall architecture to KAP-1 (31, 33). Like KAP-1, TIF1α contains an $NH_2$-terminal RBCC motif and carboxy-terminal PHD and bromodomains and can be considered an orthologue of KAP-1. However, there is little functional cross talk between these proteins: KAP-1 does not bind to nuclear hormone receptors, and TIF1α binds very weakly to the KRAB domain. When TIF1α was used in a two-hybrid screen, two of the interacting components were the murine heterochromatin proteins M31 (mMOD1) and mHP1α. Remarkably, a second two-hybrid screen using mHP1α as the bait yielded the murine homologue of KAP-1 (designated TIF1β [30]). Taken together, these data suggest that the KAP-1-mediated repression pathway may involve the local heterochromatinization of DNA templates via interaction with specific heterochromatin proteins.

A long history of studies have shown that heterochromatin is a repressive chromosomal environment (9). For example, when a euchromatic region is juxtaposed to heterochromatin by chromosomal rearrangement, the genes contained within the region become repressed. This gene-specific repression gives rise to phenotypic variegation in tissues where the genes are normally active. This phenomenon, called position effect variegation (PEV) (48), has allowed geneticists to identify second-site mutations that can modify variegation. One of the first modifiers identified at the molecular level, and subsequently the best studied, is *Drosophila* heterochromatin-associated protein 1 (HP1) (12, 13, 25). The HP1 gene, allelic to *Su(var)2-5*, is a dosage-dependent modifier of variegation (13), and the protein is diagnostic for heterochromatin (25). The exquisite sensitivity of PEV to changes in the dosage of heterochromatin proteins like HP1 has led to a model whereby heterochromatin may be envisaged as a large macromolecular complex whose constituent components are encoded by modifier genes: it is this complex that is thought to repress gene activity (50). HP1 shares a highly conserved 50- to 60-aa region, termed the chromodomain (39, 47), with another protein, Polycomb (Pc), which is a repressor of the homeotic genes (38, 39). This observation not only suggested that the chromatin-mediated silencing in PEV and repression of the homeotic genes may be mechanistically related (19, 38) but also allowed the identification of chromobox sequences in a variety of animal and plant species and the cloning of genes that are either HP1-like or Pc-like (47).

The HP1 class of chromodomain proteins are characterized by the presence of a chromodomain at the amino terminus preceded by a stretch of glutamic acid residues. Members of this class also share a second conserved domain at the carboxy terminus termed the chromoshadow domain (1). Pc-like proteins are larger and have instead of a chromoshadow domain another carboxy-terminal homology called a Pc domain (40). Three HP1-like genes have now been identified in humans and rodents. In humans, the genes have been termed hHP1α, hHP1β, and hHP1γ (18, 37, 44, 47, 57); in mice, the homologous genes have been termed mHP1α, M31 (mMOD1), and M32 (mMOD2) (22, 23, 32, 40), respectively. The characterization of M31 and M32 has been revealing. M31, which is identical to hHP1β (37), is the closest sequence homologue of *Drosophila* HP1 and is a component of constitutive heterochromatin in mice and humans (18, 37, 56). M32, the homologue of hHP1γ (57), is also a member of the HP1 class of chromobox genes (23) but is excluded from constitutive heterochromatin and is distributed in a fine-grain, or speckled, pattern of many hundreds of spots throughout the nucleoplasm. This distribution suggests that the M32 gene product is a component of a macromolecular complex that represses gene activity in euchromatic DNA through regional compaction of chromatin into a heterochromatin-like complex (23).

Our present study builds on the convergent findings from studies on transcriptional corepressors (17, 30) and from the work on mammalian HP1-like genes (23, 45). We now describe a detailed structure-function analysis in vitro and in vivo of the KAP-1 interaction with HP1 family proteins. We demonstrate that KRAB-ZFPs and KAP-1 form a stable quaternary complex with DNA and HP1 protein and that the KAP-1 interaction with HP1-like proteins occurs through a specific protein domain called the HP1 binding domain (HP1BD). We demonstrate that the M31 and M32 proteins colocalize with KAP-1 within interphase nuclei and that the location patterns of these proteins indicate that their subnuclear distribution within the nucleus is dynamic and may lead to the formation of discernible regions that may represent locally silenced chromosomal domains.

## MATERIALS AND METHODS

**Expression plasmids.** Glutathione *S*-transferase (GST) fusions of the entire murine M31 and M32 cDNAs (22) were created by subcloning into the *Eco*RI sites of pGEX-2T and pGEX-3x (Pharmacia), respectively. The hHP1α and hHP1γ cDNAs were subcloned into pGEX-2T and were kindly provided by H. J. Worman (57). GST-KRAB(B) and GST-KRAB(DV) have been described previously (17). The plasmid expressing aa 1 to 90 of KRAB fused to a $His_6$-tagged GAL4 DNA binding domain [6HisGAL4-KRAB (1-90) protein] was constructed via PCR using plasmid pM1-KOX, 1-90 (34) as a template. Briefly, a 5′ oligonucleotide incorporated a *Bam*HI site immediately 5′ to the GAL4 initiator methionine and a 3′ oligonucleotide incorporated a stop codon after aa 90 of KOX-1 followed by a *Hin*dIII site. The resulting, appropriately digested PCR product was cloned into the pQE30 vector (Qiagen Inc.) at the corresponding restriction sites. The protein was purified under denaturing conditions (6 M guanidine-HCl) and then subjected to exhaustive step dialysis. The GAL4–KAP-1 expression construct was described previously (17). The Mut1 and Mut2

GAL4–KAP-1 plasmids were created by standard PCR-mediated mutagenesis. The mutagenic primers contained the following codons: Mut1, GCTGCT (AlaAla) at amino acid positions 519 to 520; and Mut2, GAAGAG (GluGlu) at amino acid positions 487 to 488. To generate the corresponding *Escherichia coli* expression plasmids for these mutants, each was digested with *Bam*HI and *Xma*I (internal sites in human KAP-1), and the DNA fragments (encoding aa 381 to 618) were subcloned into the pQE31 (Qiagen) expression plasmid at the corresponding restriction sites. These proteins were produced, purified, and eluted from the $Ni^{2+}$-agarose (Qiagen) with imidazole under native conditions as recommended by the manufacturer. The FLAG epitope-tagged mammalian expression plasmids containing aa 1 to 191 of hHP1α and aa 17 to 173 of hHP1γ were generated by subcloning *Bam*HI/*Xho*I fragments from the pBTF4 plasmids into the corresponding sites of pcDNA3 (Invitrogen), kindly provided by H. J. Worman (57). The cytomegalovirus (CMV)-based mammalian expression plasmids used in COS-1 cells, 6HisKAP-1 delRB (aa 239 to 835), 6HisRBCC (aa 20 to 419), and 6HisPHD/Bromo (aa 619 to 835), have been described elsewhere (17, 42). The 6HisKAP-1 aa 423 to 584 mammalian expression plasmid was generated by subcloning an *Eco*RI/*Hin*dIII fragment from pQE30 (17) into the corresponding sites of pcDNA3.1 (Invitrogen). The GAL4–nuclear hormone receptor corepressor (N-CoR) plasmid expressing aa 1 to 312 of N-CoR repression domain 1 was kindly provided by M. Lazar. All PCR-derived plasmids were subjected to automated DNA sequencing of both strands to confirm sequence integrity.

**Cell extract preparation.** COS-1 cells were grown in Dulbecco modified Eagle medium (DMEM) containing 10% fetal bovine serum (FBS) and grown in 5% $CO_2$ at 37°C. Whole-cell extracts were prepared by first washing cells with phosphate-buffered saline (PBS) four times and then lysing them in the dish in ELB buffer (50 mM HEPES [pH 7.5], 250 mM NaCl, 0.1% Nonidet P-40 [NP-40], 1 mM EDTA) including protease inhibitors (0.1 mM phenylmethylsulfonyl fluoride [PMSF], 100 μg of aprotinin, 10 μg of leupeptin, and 10 μg of pepstatin per ml, and 1 mM benzamidine). The lysates were aspirated from the dish, particulate matter was clarified from the extract by centrifugation at 100,000 × *g* for 30 min, at 4°C, and supernatants were collected. COS-1 nuclear extracts (CNE) were prepared from COS-1 cells via a slight modification of the method of Lassar et al. (29). After four washes with PBS, cellular lysis was carried out by first treating cells at 4°C with nonnuclear lysis buffer (10 mM HEPES [pH 7.6], 10 mM NaCl, 1.5 mM $MgCl_2$, 20% glycerol, 0.2 mM EDTA, 0.1% Triton X-100) including protease inhibitors (0.1 mM PMSF, 100 μg of aprotinin, 10 μg of leupeptin, and 10 μg of pepstatin per ml, and 1 mM benzamidine). The nuclei were then collected by centrifugation at 1,250 × *g* for 5 min at 4°C. The pelleted nuclei were lysed in nuclear extraction buffer (NEB; 10 mM HEPES [pH 7.6], 500 mM NaCl, 1.5 mM $MgCl_2$, 20% glycerol, 0.2 mM EDTA, 0.1% Triton X-100) containing protease inhibitors. Preparations consisted of approximately $2 \times 10^7$ nuclei/ml. The extraction was carried out by rotation at 4°C for 1 h, and the tubes were then centrifuged at 100,000 × *g* for 30 min at 4°C. The final protein concentration in each nuclear extract varied from 1 to 5 mg/ml.

**Metabolic labeling.** For metabolic labeling of COS-1 cells, fresh cultures were first starved by incubation in DMEM lacking methionine and cysteine (ICN Biochemicals) for 30 min. The cells were then labeled using Tran$^{35}$S-label (75% [$^{35}$S]methionine, 15% [$^{35}$S]cysteine; ICN) for 30 to 120 min in DMEM containing 10% dialyzed FBS (Sigma). The cell cultures were then washed four times in PBS, and lysates were prepared as whole-cell or nuclear extracts as described above.

**GST protein preparation.** Following transformation of the expression plasmids into competent *E. coli* BL21(DE3) bacteria and identification of highly expressing bacterial colonies, 10-ml overnight cultures were started in 2YT medium. The next day, the entire overnight culture was added to 250 ml of fresh 2YT, and the culture was allowed to grow until the optical density at 600 nm reached 0.4 to 0.6. Isopropyl-β-D-thiogalactopyranoside was then added to 0.5 mM, and the cultures were incubated for an additional 3 to 4 h. The cells were pelleted at 8,000 × *g* for 10 min at 4°C. The bacteria were resuspended in 4 ml of PBS, and 400 μg of lysozyme was added. After a 15-min incubation on ice, dithiothreitol to 5 mM and protease inhibitors to final concentrations of 0.1 mM PMSF, 100 μg of aprotinin per ml, 10 μg of leupeptin per ml, 10 μg of pepstatin per ml, and 1 mM benzamidine were added. Sarcosyl was added to a final concentration of 3.5%, and the bacterial suspension was sonicated for 30 s, left on ice for 1 min, and then sonicated for an additional 30 s. The sample was centrifuged 16,000 × *g* for 10 min at 4°C. Triton X-100 was then added to the supernatant to a final concentration of 4%, and the protein extract was snap frozen in small aliquots. Large-scale preparations of purified GST fusion proteins were prepared by eluting the proteins from glutathione-Sepharose (Pharmacia) in GST elution buffer (100 mM Tris [pH 8.0], 150 mM NaCl, 0.1% NP-40, 20 mM freshly added reduced glutathione). Elution was with a buffer volume equal to 2.5 times the packed bead volume, and incubation was at room temperature for 1 h. The beads were centrifuged briefly, and the supernatant was collected. The elution was repeated, and the supernatants were combined and concentrated in 5,000-molecular-weight-cutoff microspin concentrators (Millipore). Protein concentrations were determined for each of the proteins by the DC protein assay (Bio-Rad). The GST proteins were diluted with PBS prior to use in all electrophoretic mobility shift assays (EMSAs).

**GST pull-down assays.** Five micrograms of freshly prepared GST fusion protein immobilized on glutathione-Sepharose (Pharmacia) was incubated with either 2 μl of in vitro-translated, $^{35}$S-labeled KAP-1 (T3 TnT; Promega), 50 μl of $^{35}$S-labeled whole-cell lysate from transiently transfected COS-1 cells, 500 μg of HeLa whole-cell lysate (PBS, 0.1% NP-40), or 1 to 2 μg of $Ni^{2+}$-agarose (Qiagen)-purified recombinant $His_6$-tagged protein in 200 μl of BB100 (20 mM Tris [pH 7.9], 100 mM NaCl, 0.2 mM EDTA, 10% glycerol, 0.1% NP-40, 1 mM PMSF, 500 μg of bovine serum albumin [BSA; fraction V]) for 1 h at room temperature. Protein complexes were washed four times with BB750 (20 mM Tris [pH 7.9], 750 mM NaCl, 0.2 mM EDTA, 10% glycerol, 0.1% NP-40, 1 mM PMSF), and the bound proteins were eluted in 2× Laemmli buffer by boiling for 10 min. Proteins were resolved by standard sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) procedures. Retention of $^{35}$S-labeled KAP-1 was visualized by fluorography in dimethyl sulfoxide–2,5-diphenyloxazole (Fisher Biotech) followed by exposure to MR X-ray film (Kodak). For evaluation of KAP-1 binding to GST fusion protein by Western blotting, proteins were transferred to Immobilon-P (Millipore) in Towbin buffer–0.1% SDS for 20 h at 4°C and 250 mA. Membranes were blocked in 5% Blotto–Tris-buffered saline (TBS; 50 mM Tris [pH 7.5], 150 mM NaCl). KAP-1 was detected with antigen affinity-purified rabbit polyclonal antiserum (17) diluted 1:200 in TBS–1% BSA. Membranes were washed three times in TBS–0.05% Tween 20 and then incubated with a 1:5,000 dilution of a horseradish peroxidase-conjugated goat anti-rabbit immunoglobulin G secondary antibody (Sigma) in 1% Blotto-TBS. Proteins were visualized by chemiluminescence (Pierce) and exposure to MR X-ray film (Kodak).

**EMSAs.** The ability of KAP-1 to bind a KRAB domain-containing protein was assayed by EMSA essentially as described previously (16). Two partially overlapping GAL4-upstream activation sequence (UAS) oligonucleotides (5′-GAT CCCGGAGGACAGTACTC-3′ and 5′-CTAGACGGAGTACTGTCCTC-3′) were annealed and used for EMSAs. EMSA binding reaction mixtures (20 μl) were assembled by adding 50 ng of purified recombinant 6HisGAL4-KRAB to 2× binding buffer (40 mM HEPES [pH 7.6], 100 mM NaCl, 1 mM dithiothreitol, 10 mM $MgCl_2$, 20% glycerol) containing 1 μg of poly(dI-dC) · poly(dI-dC). Nuclear extract (~5 μg/l to 2 μl) and/or titrating amounts of GST fusion protein were added, and incubation was continued at room temperature for 15 min. Then γ-$^{32}$P-end-labeled GAL4-UAS probe was added (~0.5 to 1 ng, $10^5$ cpm/reaction), and the incubation continued for an additional 10 min at room temperature. The samples were chilled on ice, centrifuged 4 min at 4°C, and then loaded onto 4.5% nondenaturing polyacrylamide gels containing 1× TBE (45 mM Tris [pH 8.3], 45 mM boric acid, 1 mM EDTA), which were preelectrophoresed for 45 min at 4°C in 0.5× TBE buffer. DNA-protein complexes were separated from unbound DNA by electrophoresis at 500 V and 24 mA for 2 more h at 4°C. The gels were dried and exposed to MR X-ray film.

**FPLC gel filtration fractionation.** The $^{35}$S-labeled CNE was subjected to isocratic fast protein liquid chromatography (FPLC) gel filtration on a Superose 6 column (Pharmacia) in NEB running buffer at a flow rate of 0.5 ml/min at room temperature. The KAP-1 protein was detected in the FPLC fractions by immunoprecipitation using protein G-purified KAP-1 polyclonal antibodies (17). To determine if a KAP-1–heterochromatin protein complex is stable to and detectable after FPLC gel filtration, a nonradioactive CNE containing KAP-1 was incubated with the purified GST-heterochromatin fusion protein, GST-M32, in NEB. After 60 min, incubation at room temperature, the mixture was subjected to isocratic gel filtration as described above. Glutathione-Sepharose resin (20 μl, 50% slurry; Pharmacia) was added to individual FPLC fractions, incubated for 1 h at 4°C followed by 1 h at room temperature, and then washed two times with NEB and five times with PBS; 2× Laemmli sample buffer was added to the resin; samples were boiled and separated by SDS-PAGE (10% gel). The proteins were transferred to Immobilon-P membranes, and recovered KAP-1 was visualized by using protein G-purified anti-KAP-1 polyclonal antibodies in a standard Western blot procedure described above.

**Indirect immunofluorescence.** NIH 3T3 cells were grown on glass coverslips in DMEM containing 10% calf serum and immunostained as previously described (35). The murine KAP-1 protein was visualized by indirect immunofluorescence with an antigen affinity-purified rabbit polyclonal antibody previously described (42). The M31 protein was visualized by indirect immunofluorescence using a rat monoclonal antibody (MAb) raised to the COOH-terminal 71 amino acids (anti-M31 MAb MAC 353 [56]). The M32 protein was detected with a rat MAb developed by using a GST fusion protein that included the entire coding region of M32 (anti-M32 MAb MAC 385 [23]). The hHP1 proteins were recognized by using a rabbit polyclonal antibody raised against hHP1, kindly provided by W. C. Earnshaw (44). The secondary antibodies were either Texas red-conjugated goat anti-rabbit or biotinylated goat anti-rat, used in conjunction with an avidin-biotin-linked fluorescein isothiocyanate reagent (Vector Laboratories). All immunofluorescence was performed as described previously (35). DNA was counterstained with Hoechst 33258 (Sigma), and coverslips were mounted with Fluoromount G (Fisher Scientific). Cells were visualized with a scanning confocal microscope (Leica Inc.). The images obtained through image capture were processed with Adobe Photoshop 3.0.4 (Adobe Systems Inc.) from files or scanned slide images.

**Transient transfection luciferase assays.** DNA for transfection was prepared by CsCl gradient centrifugation. Protein expression from all plasmids was confirmed by transient transfection of COS-1 cells followed by immunoprecipitation of [$^{35}$S]methionine-labeled cell extracts as described previously (16). All transcription assay transfections were done with NIH 3T3 cells maintained in DMEM–10% calf serum. A total of $2.0 \times 10^5$ cells were plated in a 60-mm-

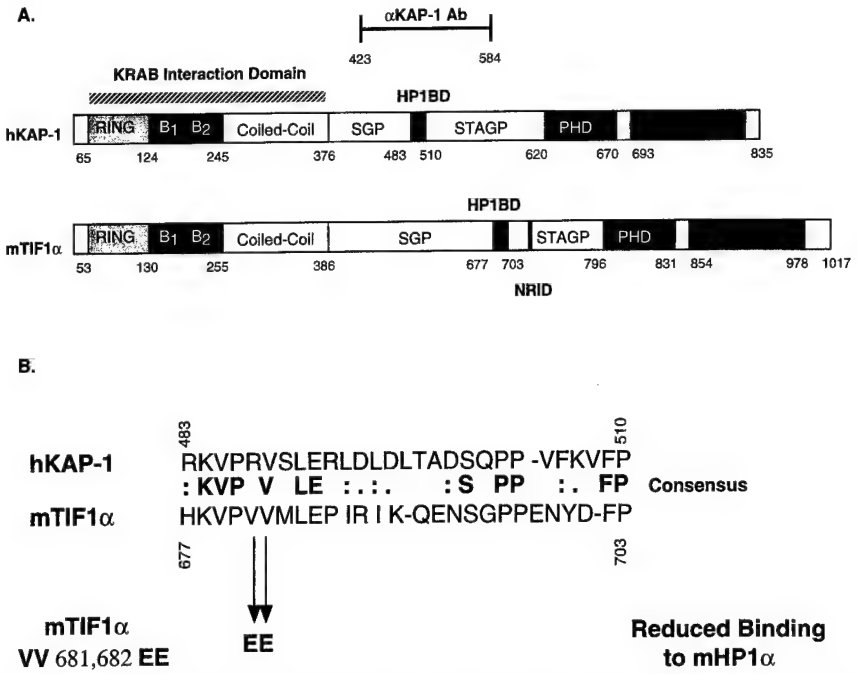



FIG. 1. Schematic diagram illustrating the architecture of the KAP-1/TIF1α family of transcriptional regulatory proteins. (A) The conserved motifs include the RING finger, B boxes ($B_1$ and $B_2$), coiled-coil, PHD (also known as the LAP domain), and bromodomain (Bromo). Note the overall similar architectures among this family of proteins defined by the layout of the various domains. The putative HP1BD (black boxes) is conserved in each protein; the nuclear receptor interaction domain (NRID) is conserved only in TIF1α. Regions of significantly enriched amino acids, serine-glycine-proline (SGP) and serine-threonine-alanine-glycine-proline (STAGP), are also spatially conserved in this family. The minimal KRAB binding domain comprises the RBCC domain and is marked with a striped bar above the hKAP-1 protein. Also indicated is the region of hKAP-1 expressed as a recombinant protein and used to raise anti-KAP1 polyclonal antisera (αKAP-1 Ab). (B) Amino acid alignment of the putative HP1BD of KAP-1 and mTIF1α. Mutation of valines 681 and 682 of mTIF1α to glutamic acid (VV 681,682 EE) were previously observed to abolish mHP1α binding to mTIF1α (30). The corresponding sequences are from data bank entries 78773 (hKAP-1) and 78219 (mTIF1α).

diameter tissue culture dish and transfected in OptiMEM for 5 to 6 h with Lipofectamine (Life Technologies Inc.) under conditions recommended by the manufacturer. The cells were harvested 24 h posttransfection, and luciferase assays were performed as previously described (17). Cotransfection with a pcDNA3-β-galactosidase expression plasmid was used to normalize all luciferase values.

## RESULTS

**Sequence analysis of KRAB binding corepressor and heterochromatin-associated protein families.** TIF1α and KAP-1 share a number of similar amino acid sequence motifs, including the well-characterized RING finger, $B_1$ and $B_2$ boxes, a coiled-coil, a PHD finger, and an extended bromodomain (17, 30, 36) (Fig. 1A). Overall, KAP-1 is only 33% identical and 45% similar to hTIF1α. Previous studies with mTIF1α delineated a region of the protein (aa 672 to 698) which may be sufficient for binding mHP1α (30) (Fig. 1A). We have completed sequence comparison of these amino acids to KAP-1 and have found that KAP-1 contains a highly homologous domain which we believe is a potential HP1BD (KAP-1 aa 483 to 510). This putative HP1BD in KAP-1 is conserved with 45% identity and 60% similarity to the analogous region in mTIF1α (Fig. 1B). The greatest degree of identity and conservation is shown in the amino-terminal portion of the HP1BD. Two previously identified valines (aa 681 and 682) were mutated to glutamic acid in mTIF1α and shown to substantially reduce binding to mHP1α (30). One of these valines is among the conserved amino acids found in this putative HP1BD (Fig. 1B).

Sequence comparisons among the human and mouse chromodomain-containing HP1 proteins suggest that the murine homologues of hHP1α, -β, and -γ are mHP1α, M31, and M32, respectively. Studies have shown that all of these HP1 proteins contain the highly conserved chromodomains and chromo shadow domains (1) as well as nuclear localization signals, and some may even have potential DNA binding domains (49). To evaluate antibodies raised against the M31, M32, and hHP1 proteins, we measured the binding specificity and cross-reactivity of these antibodies against bacterially expressed and affinity-purified human and murine heterochromatin fusion proteins. The results indicate that the M31 antibody is specific for M31, the M32 antibody recognizes both M32 and hHP1γ, and the hHP1 antibody recognizes all of the heterochromatin proteins (Fig. 2). These observations suggest that the M31 and M32 antibodies are specific reagents that can be used in immunolocalization studies with KAP-1.

**KAP-1 and HP1 family proteins interact in vitro.** In light of the result that KAP-1 could bind mHP1α and M31 proteins in a yeast two-hybrid assay (30), we initiated a comprehensive analysis of the abilities of other heterochromatin family members to bind KAP-1. After purification of GST-heterochromatin protein fusion proteins (Fig. 3A), the resins were analyzed for the ability to bind KAP-1 protein produced by in vitro transcription and translation. In general, significant binding of KAP-1 was observed for all of the HP1 proteins but was negative for the control GST protein (Fig. 3B). Further analysis revealed that the interaction between in vitro-transcribed-translated KAP-1 and recombinant HP1 proteins in solution is extremely stable, as nearly equal signal intensities were observed between resins washed in either 250 mM NaCl or 1 M NaCl (data not shown). To assay the ability of cell-derived KAP-1 protein to bind HP1 proteins, GST-hHP1α was incubated with HeLa cell extracts. Binding of KAP-1 to the recombinant GST-hHP1α protein was detected by Western blot analysis using polyclonal antiserum raised against KAP-1. As

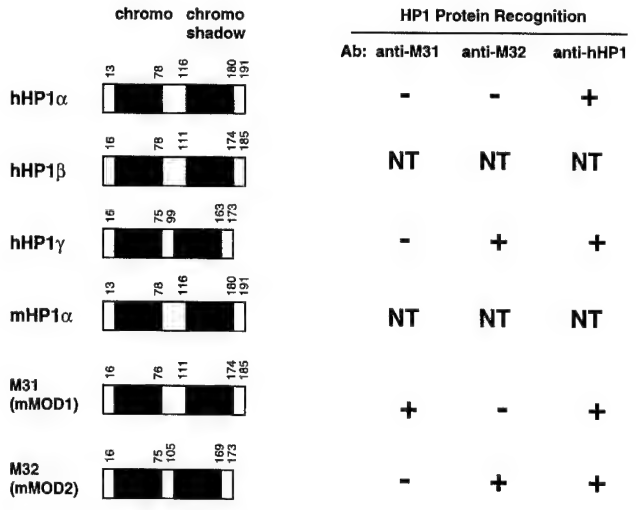



FIG. 2. The HP1 family of human and mouse heterochromatin proteins. The regions of conservation which are diagnostic for the HP1 family of proteins include the $NH_2$-terminal chromodomain (black boxes) and the COOH-terminal chromo shadow domain (grey boxes). On the right is a summary of the immunoreactivities for the various antibodies (Ab) available to recombinant heterochromatin proteins derived from humans and mice. NT, not tested. The corresponding sequences are from the following data bank entries: hHP1α, 60277; hHP1β, 23197; hHP1γ, 26312; mHP1α, 99641; M31, 56690; and M32, 56683.

illustrated in Fig. 3C, hHP1α displays significant binding capacity for cell-derived KAP-1; greater than 50% of the input KAP-1 in the extract bound to the GST-hHP1α resin. Identical data were obtained when GST-hHP1γ resin was used in the assays (data not shown). The GST-KRAB and GST-KRAB(DV) resins demonstrate the specificity of KAP-1 binding, since KAP-1 is observed to interact only with GST-KRAB and not the mutant GST-KRAB(DV), which lacks KAP-1 binding activity (Fig. 3C). To grossly localize the HP1 interaction region, we expressed truncated KAP-1 proteins in COS-1 cells and assayed extracts containing these proteins for the ability to bind GST-hHP1α (data not shown). As summarized in Fig. 4, only KAP-1 proteins which contained the central region spanning aa 423 to 584, which includes the putative HP1BD, bound to GST-hHP1α. In addition, two peptides corresponding to aa 435 to 449 and 568 to 581 of KAP-1 when added in molar excess to the binding reaction mixture were unable to block the interaction between KAP-1 and HP1s, suggesting that the putative HP1BD is between aa 450 and 568 of KAP-1 (data not shown). To determine if the KAP-1–HP1 interaction is direct, we used purified, recombinant proteins (Fig. 5). The region of wild-type KAP-1 spanning aa 381 to 618 and two mutant versions thereof (Fig. 5A) were expressed in bacteria and purified to homogeneity (Fig. 5B). These proteins were used in in vitro binding assays utilizing GST-hHP1α and -γ proteins purified to near homogeneity (Fig. 3A). Wild-type KAP-1 efficiently bound to both resins (Fig. 5C). The Mut1 protein, which contains the LI-AA (aa 519 and 520) substitutions carboxy terminal to the putative HP1BD, also bound each resin with moderate efficiency. However, the Mut2 protein, with the RV-EE (aa 487 and 488) substitutions within the putative HP1BD, completely abolished interaction with the GST-hHP1 proteins. Together, these results strongly suggest that the KAP-1–HP1 association is via direct protein-protein interactions and that the region of KAP-1 spanning aa 450 to 568 possesses the interaction domain.

Previously, we have shown that KAP-1 can interact very efficiently with KRAB domain-containing proteins in an EMSA (17, 42). This highly sensitive assay allows for the detection of protein-protein interactions with limiting amounts of proteins at concentrations at or close to their dissociation constants, rather than at extreme excess, as is usually the case in solution-based GST binding assays. In addition, the EMSA requires that the complex once formed be stable to prolonged electrophoresis and has the capability of distinguishing high- and low-affinity interactions. To further assess the ability of the different HP1 proteins to interact with KAP-1, we again used EMSA. The well-characterized GAL4 (aa 1 to 147) DNA binding domain was used to construct a 6HisGAL4-KRAB fusion protein, which was purified to near homogeneity. This protein binds with moderate efficiency to the $^{32}P$-labeled GAL4-UAS oligonucleotide (Fig. 6) under EMSA conditions. However, preincubation of the 6HisGAL4-KRAB with a KAP-1-containing CNE results in a dramatic supershift of the DNA binding complex in the EMSA gel. This supershifted complex contains both proteins, as antibodies directed to either GAL4 or KAP-1 efficiently supershifted this complex further (data not shown). Importantly, purified, recombinant HP1 fusion proteins GST-M31, GST-M32, GST-hHP1α, and GST-hHP1γ were all able to bind the 6HisGAL4-KRAB–KAP-1 complex, as evidenced by a new supershifted complex containing the HP1 protein. hHP1α appears to have a lower affinity for the DNA–KRAB–

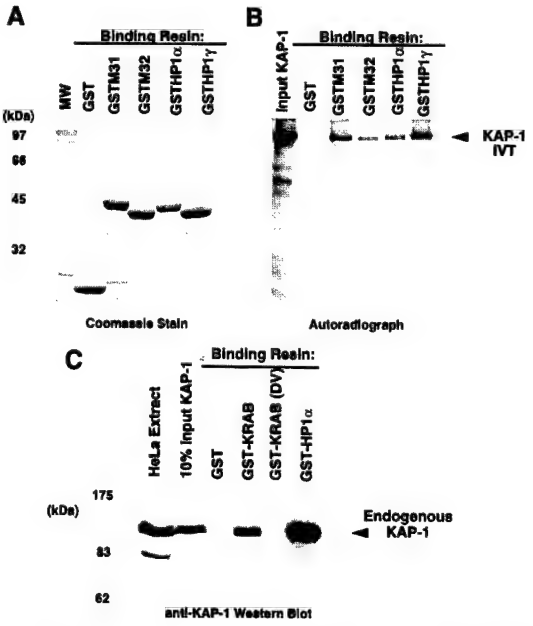



FIG. 3. The KAP-1 protein binds to both human and murine HP1 family proteins. (A) Analysis of the GST fusion proteins affinity purified as a soluble protein from *E. coli* extracts and used in the binding assays. Approximately 5 μg of each protein was electrophoresed on an SDS–10% polyacrylamide gel and stained with Coomassie blue. Equivalent quantities of protein (5 μg) were used for each binding reaction as determined by Coomassie blue staining and quantification against BSA. MW, molecular weight markers. (B) hKAP-1 was prepared by in vitro transcription-translation (IVT) of a full-length human cDNA, and the [$^{35}S$]methionine-labeled protein (arrow) was used in binding reactions with the indicated GST fusion proteins. The input lane represents the total amount of KAP-1 in each reaction mixture. No binding was observed in reactions containing GST alone. (C) A HeLa whole-cell extract containing endogenous KAP-1 was used in binding reactions with the indicated GST fusion protein. GST-KRAB represents a positive control for KAP-1 interaction, and GST-KRAB(DV) is a mutant with reduced affinity for KAP-1, which served as a negative control for KAP-1 interaction. Retention of KAP-1 by the GST fusion protein was detected by Western blot analysis using affinity-purified anti-KAP-1 antibodies. The input lane represents 10% of the KAP-1 in each binding reaction as detected by Western blot analysis of the extract.

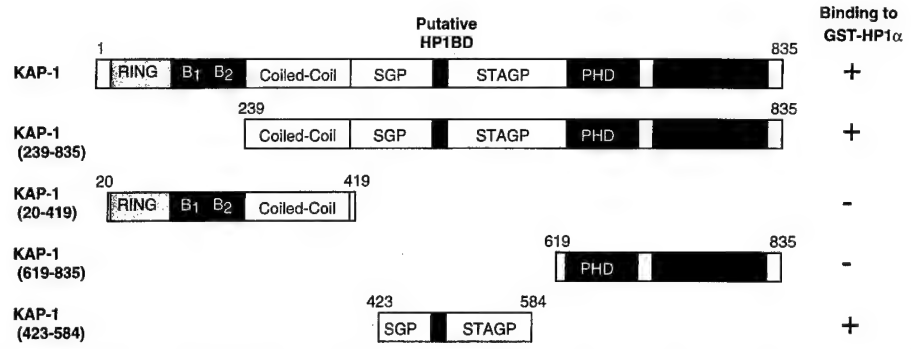



FIG. 4. Localization of the HP1BD in KAP-1. COS-1 cells were transfected with the indicated expression constructs containing different regions of KAP-1, and extracts from these cells were assayed for the ability of KAP-1 to bind GST-hHP1α. All KAP-1 molecules which included the putative HP1BD bound to GST-hHP1α, as indicated by the plus signs. For notation, see the legend to Fig. 1.

KAP-1 ternary complex (Fig. 6). The addition of purified GST protein alone failed to produce the supershift. We also tested all of the heterochromatin proteins for binding by varying the order of addition of each component (data not shown). In these experiments, the GAL4-UAS–6HisGAL4-KRAB–KAP-1 complex was formed prior to addition of the heterochromatin protein: the binding observed was identical to that seen in Fig. 6, suggesting that interaction of KRAB and KAP-1 does not preclude binding of the heterochromatin protein. These results strongly suggest that a quaternary GAL4-UAS–6HisGAL4-KRAB–KAP-1–HP1 complex is readily formed in vitro and is stable to the EMSA conditions used.

Using gel filtration chromatography, we have shown that the endogenous KAP-1-containing complex found in CNE migrated at approximately 560 kDa (42). We sought to determine if an exogenous source of HP1 protein could influence the migration of the CNE KAP-1 complex in gel filtration (Fig. 7). Purified soluble GST-M32 protein was incubated with CNE, and then the mixture was subjected to gel filtration. The fractions were incubated with glutathione-Sepharose beads to recover the GST-M32 protein and any associated proteins, and the washed beads were analyzed for KAP-1 by Western blotting. Comparison of the native 560-kDa KAP-1 complex from CNE to that produced when GST-M32 was mixed with CNE prior to separation shows that the KAP-1 from the nuclear extract was efficiently bound to and recovered by binding GST-M32 after separation (Fig. 7). More importantly, the resulting KAP-1 complex now migrated at approximately 800 kDa. These results demonstrate that the reconstituted KAP-1–M32 complex is stable to gel filtration and suggest that multiple molecules of M32 may bind to the endogenous KAP-1 complex.

**Colocalization of KAP-1 with the M31 and M32 HP1-like proteins.** To ascertain a potential in vivo role for the KAP-1–HP1 association, we used antibodies raised to M31 (56) and M32 (23) concurrently with an affinity-purified anti-KAP-1 antibody in dual staining-indirect immunofluorescence studies of asynchronous NIH 3T3 fibroblasts. In a random field of immunostained cells, there appeared to be various levels of KAP-1 staining (Fig. 8 and 9). In general, KAP-1 staining was found exclusively throughout the nucleoplasm as an evenly distributed grainy (speckled) pattern which was almost always excluded from nucleoli. In nearly 50% of the NIH 3T3 nuclei, KAP-1 is significantly concentrated in dot-like structures which could represent regions of pericentromeric heterochromatin commonly observed in murine cells (Fig. 9A). It should be noted that this specific staining pattern for KAP-1 does not overlap with other well-characterized nuclear dot-like structures previously described for PML, SC35, and BRCA1 (34a). KAP-1 staining in human HEp-2 cells yielded very similar results (Fig. 9D to F). An interesting pattern was observed in less than 5% of the nuclei: KAP-1 was concentrated both at the

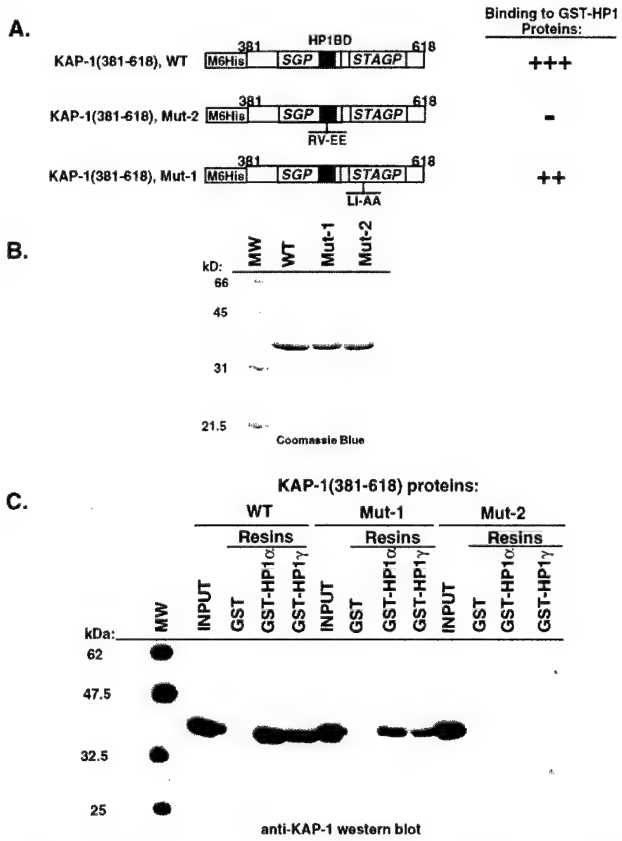



FIG. 5. Amino acid substitutions in the HP1BD abolish HP1 binding in vitro. (A) The indicated wild-type (WT) or mutated segment of KAP-1 (for notation, see the legend to Fig. 1) was expressed in bacteria and purified by nickel chelate chromatography under native conditions. These highly purified proteins (B) were used in GST binding assays using the recombinant proteins shown in Fig. 3A. After the binding reactions and extensive washing, the amount of bound KAP-1 protein was determined by Western blot analysis using affinity-purified anti-KAP-1 sera (C). The input lane represents the total amount of recombinant KAP-1 added to each binding reaction mixture. No binding was detected for GST alone. Note that the recombinant Mut2 protein, which possesses the RV-EE mutation in the putative HP1BD, completely abolishes binding to both hHP1α and hHP1γ. MW, molecular weight markers.

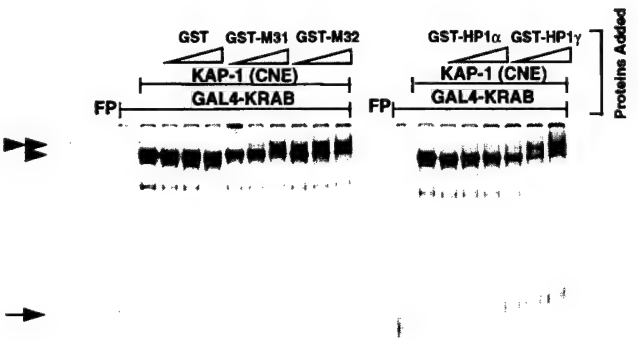



FIG. 6. Purified recombinant HP1 proteins can bind a DNA–GAL4-KRAB–KAP-1 ternary complex. A gel shift assay was performed by incubating a highly purified recombinant 6HisGAL4-KRAB domain protein with KAP-1 derived from a CNE, the indicated GST fusion protein (increasing concentrations from left to right indicated by open triangles: 50, 100, and 200 ng each of GST, GST-M31, GST-M32, GST-hHP1α, and GST-hHP1γ), and a $\gamma$-$^{32}$P-labeled GAL4-UAS oligonucleotide probe. The complexes were separated by nondenaturing PAGE, and DNA-protein interactions were detected by autoradiography. The symbols on the left denote the following complexes: arrow, DNA–6HisGAL4-KRAB; single arrowhead, DNA–6HisGAL4-KRAB–KAP-1 ternary complex; double arrowhead, DNA–6HisGAL4-KRAB–KAP-1–HP1 quaternary complex. The free probe (FP) was allowed to run off the bottom of the gel to allow for better separation of the complexes at the top of the gel.

periphery of pericentromeric heterochromatin and around nucleoli (Fig. 8D and G). Since the cell populations illustrated are asynchronously growing, log-phase cells, these data may suggest that the different subnuclear localization patterns of KAP-1 observed in a random field of cells may be due to cell cycle regulation. In support of such a conclusion, apparent daughter cells in many of the fields shown display very similar KAP-1 staining patterns.

Immunostaining for the M31 protein revealed two distinct patterns in interphase nuclei. The less frequently observed pattern was characterized by diffuse staining throughout the nucleoplasm, with little or no localization to the nucleoli or heterochromatic regions (Fig. 8B). More typically, staining of the M31 protein was restricted to large condensed heterochromatic regions (Fig. 8E and H), similar to results previously shown in interphase mouse C1271 cells (23). When the KAP-1 and M31 staining patterns were directly compared, we found that in some cells, the signals appear to be juxtaposed (Fig. 8C and F). Of particular interest was the significant colocalization of KAP-1 with M31 in the regions bordering the nucleoli and pericentromeric heterochromatin (Fig. 8G to I).

Immunostaining of the M32 protein was generally observed as a fine, speckled pattern throughout the nucleoplasm in regions consistent with euchromatin (Fig. 9B, E, and H). This particular staining pattern is consistent with previous reports for M32 (23). When M32 staining was directly compared to KAP-1 staining, we found that some of these punctate and small speckled M32-reactive regions were also domains where high concentrations of KAP-1 localized (Fig. 9C, F, and I). In some nuclei, M32 and KAP-1 were nearly exclusively colocalized (Fig. 9G to I). Since the M32 MAb also reacts with hHP1$\gamma$ (Fig. 2), we looked at the localization of HP1$\gamma$ in human HEp-2 cells. The pattern of hHP1$\gamma$ staining appeared to be very similar to that observed for M32 in NIH 3T3 cells: punctate nuclear localization in regions that appear to be consistent with euchromatin (Fig. 9D to F). Again significant colocalization of hHP1$\gamma$ and KAP-1 was observed in the merged images (Fig. 9F). This colocalization may depend on cell cycle stage, which is suggested by the similar staining patterns observed for two daughter cells (arrows) in this field (Fig. 9F). In summary, both the M31 and M32 heterochromatin proteins appear to display substantial colocalization with endogenous KAP-1 in interphase nuclei. These data imply a potential in vivo association between KAP-1 and members of the HP1 family, which may function to epigenetically silence a KRAB-ZFP-regulated gene via heterochromatinization.

**Modulation of KRAB–KAP-1 repression activity by hHP1α and the KAP-1 HP1BD.** To determine if there were any functional consequences of the KAP-1–HP1 interaction, we cotransfected NIH 3T3 fibroblasts with a GAL4-KRAB plasmid and the GAL4-UAS luciferase reporter plasmid (Fig. 10A). Under these conditions, the GAL4-KRAB protein was a strong repressor of the luciferase reporter gene activity; however, this repression was further enhanced by cotransfection of the hHP1α expression plasmid in a concentration-dependent manner (Fig. 10B). Cotransfected hHP1α also enhanced GAL4–KAP-1-mediated repression, although to a lesser extent. A GAL4 fusion of the $NH_2$-terminal repression domain of the N-CoR corepressor also displayed significant repression activity, yet this activity was not enhanced by cotransfected hHP1α (Fig. 10B). Thus, N-CoR may not utilize heterochromatin proteins for mediating repression. As a complementary strategy, we attempted to disrupt endogenous KAP-1–HP1 complexes by overexpressing a protein which encodes the segment of KAP-1 (aa 423 to 584) which contains the HP1BD. Ectopic expression of this protein would be predicted to compete with endogenous KAP-1 for binding to endogenous HP1 family proteins, resulting in relief of repression via a squelching mechanism. As predicted, increasing amounts of KAP-1

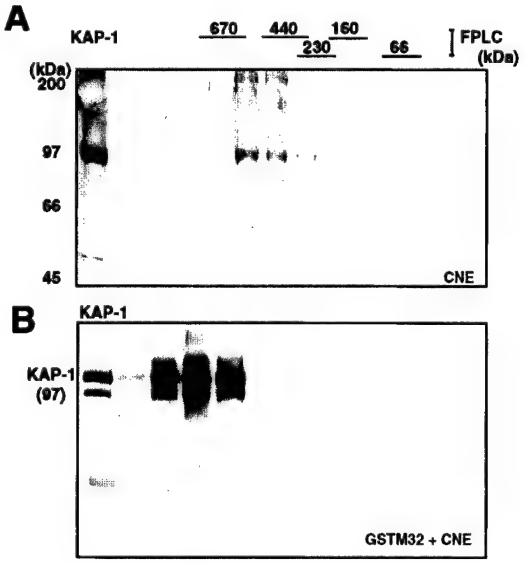



FIG. 7. Cell-derived KAP-1 complexes increase in apparent size in the presence of the heterochromatin protein M32. (A) A [$^{35}$S]methionine-labeled CNE was subjected to gel filtration (Superose 6), and the column fractions were analyzed for KAP-1 by immunoprecipitation. (B) A nonradioactive CNE was incubated with purified GST-M32 protein, and the mixture was subjected to gel filtration as for panel A. The column fractions were incubated with glutathione-Sepharose beads and washed, and the retained KAP-1 was detected in a Western blot assay using affinity-purified anti-KAP-1 sera. Lane 1 is a sample of the unfractionated CNE containing endogenous KAP-1 as a positive control and size marker. Indicated above the gels are the positions of the molecular mass standards used to calibrate the Superose 6 column, which were determined by Coomassie blue staining of duplicate gels.

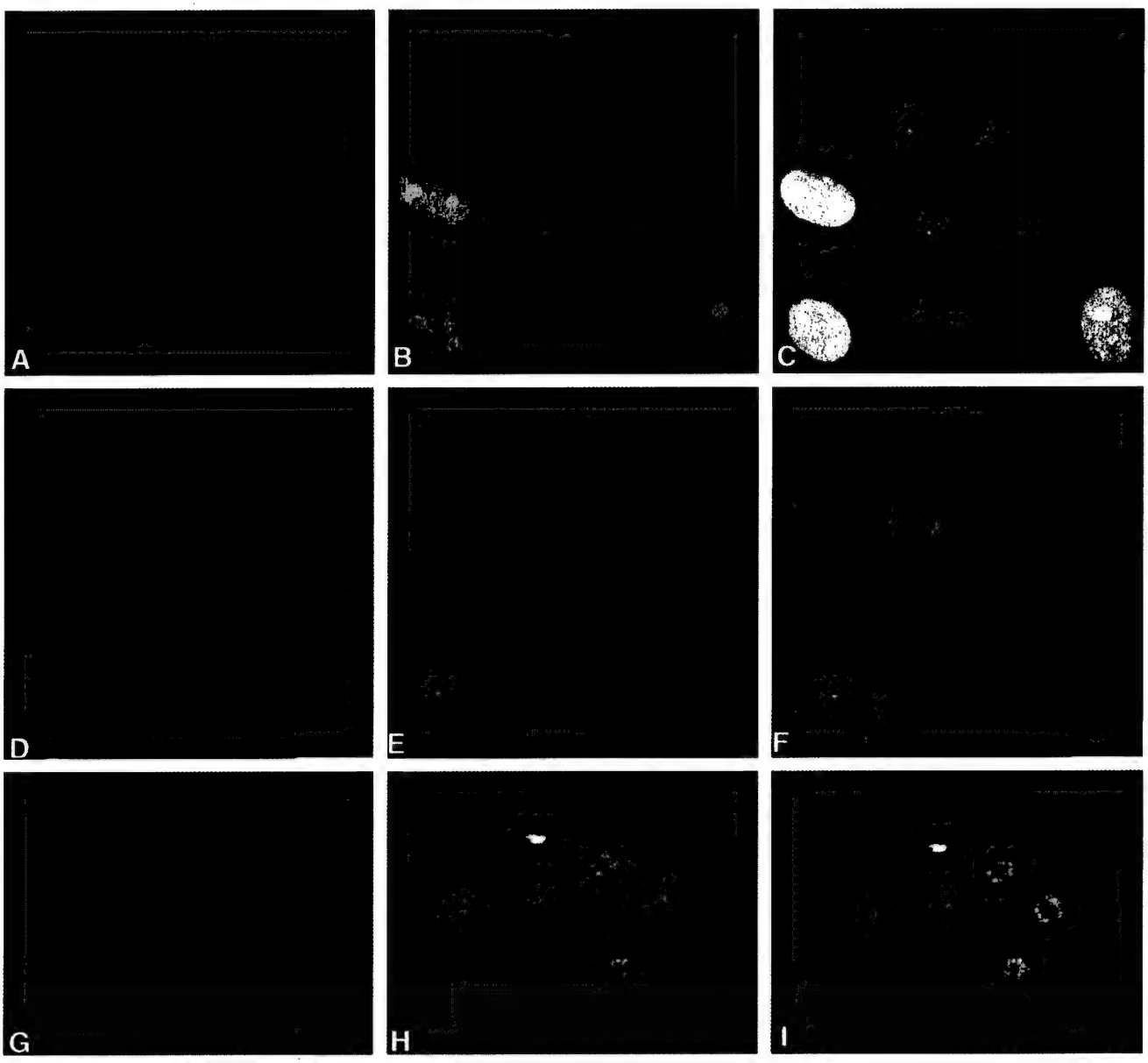



FIG. 8. Colocalization of endogenous KAP-1 with the heterochromatin protein M31 in interphase nuclei. All images were derived from the staining of a single population of cells simultaneously for both antigens and viewed by confocal microscopy. (A to F) Each row illustrates an independent random field of asynchronously growing NIH 3T3 fibroblasts to provide a representative view of the various staining patterns observed for both KAP-1 and M31. Staining of individual antigens is identified by the color given in the upper corners of each image. The merged images are shown on the far right for each set. As illustrated, staining for both antigens was exclusively nuclear, with very little background staining in the cytoplasm. KAP-1 staining demonstrated a grainy, speckled pattern throughout the nucleus which was largely excluded from nucleoli. M31 staining displayed a strong association with large heterochromatic granules. (G to I) High-magnification single-cell analysis of KAP-1 and M31 subnuclear distribution. Each panel shows a single nucleus from the field displayed in panels D to F. Note the discrete pericentromeric heterochromatin and nucleoli, as detected by M31 (56), and the concentration of KAP-1 around these structures (I).

HP1BD (aa 423 to 584) efficiently relieved GAL4-KRAB-mediated repression but had no effect on the basal level of transcription of the luciferase vector in the absence of GAL4-KRAB (Fig. 10C). We next determined the effect of mutations in KAP-1 which abolish HP1 binding (Fig. 11). The Mut1 and Mut2 mutations (Fig. 5) were introduced into the GAL4-KAP-1,293-835 expression plasmid (Fig. 11A), which has been previously shown to contain the repression functions of KAP-1. Each of these plasmids express equivalent amounts of nuclearly localized (data not shown) proteins which are readily detected by both anti-KAP-1 and anti-GAL4 sera (Fig. 11B). As expected, expression of the heterologous wild-type GAL4–KAP-1 protein demonstrated a strong, dose-dependent repression of the GAL4-UAS-simian virus 40 (SV40) immediate-early promoter luciferase reporter (Fig. 11C). The heterologous Mut2 protein, which is negative for HP1 binding in vitro (Fig. 5C), was strongly reduced in its repression capacity. Similar to its HP1 binding characteristics in vitro, expression of the heterologous Mut1 protein was less strongly affected. The residual repression activity in both Mut1 and Mut2 proteins is not surprising, as the still-intact PHD and bromodomain in each protein are themselves independent re-

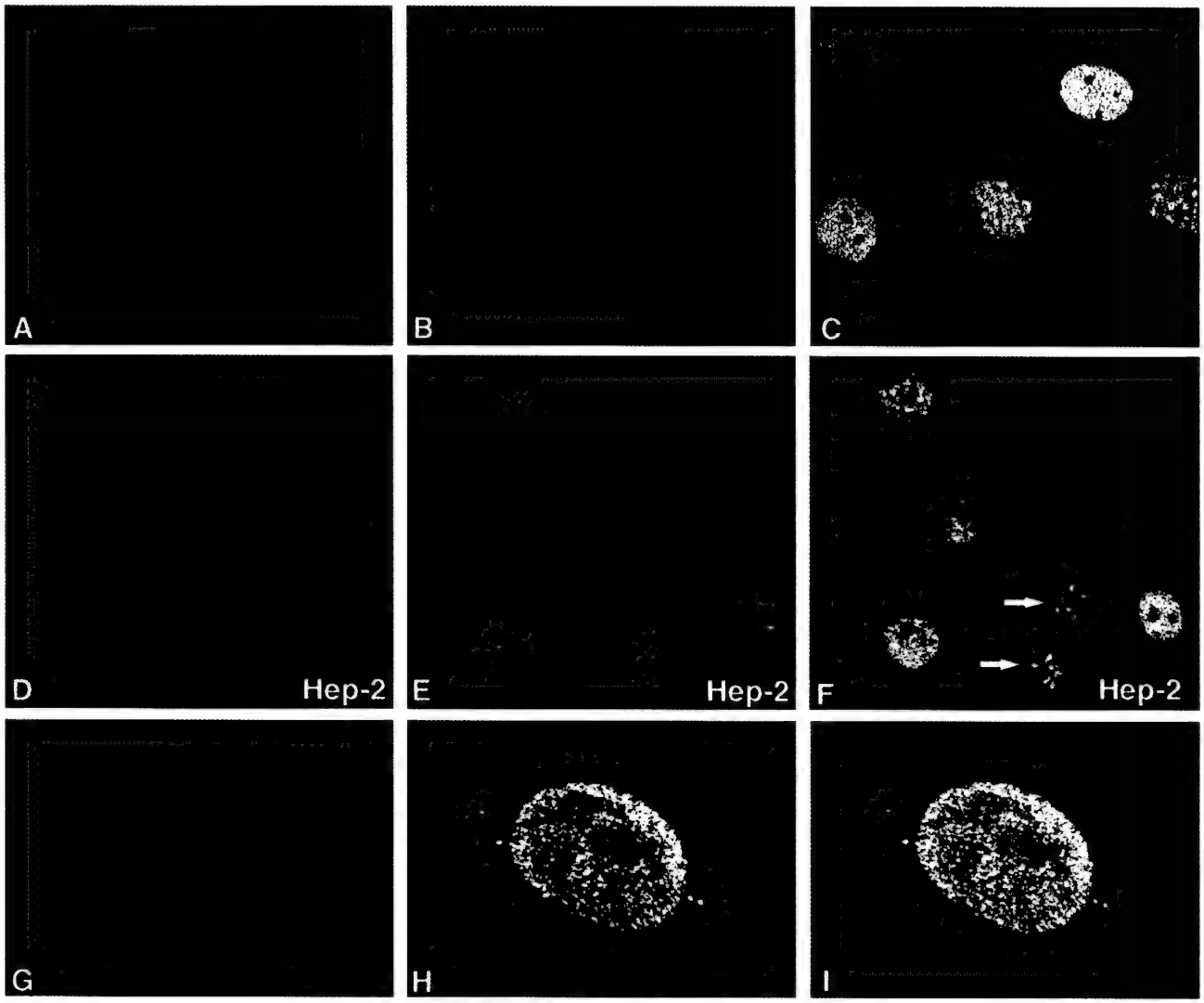



FIG. 9. Colocalization of endogenous KAP-1 with the heterochromatin protein M32 in interphase nuclei. All images were derived from the staining of a single population of cells simultaneously for both antigens and viewed by confocal microscopy. (A to C) Random fields of asynchronously growing NIH 3T3 fibroblasts. (D to F) KAP-1 and hHP1γ staining in human HEp-2 cells, as indicated in the lower right corner of each image (arrows point to two apparent daughter cells). Note the speckled pattern for M32 which is largely excluded from nucleoli and pericentromeric heterochromatin. In several cells, a significant overlap was observed between KAP-1 and M32 staining (C, F, and I). The most dramatic representation of this colocalization is portrayed in panels G to I, representing a high-magnification single-cell analysis of KAP-1 and M32 subnuclear distribution in a NIH 3T3 nucleus, where the green and red speckles show a near complete overlap, yielding a uniform yellow in the merged image (I).

pression domains (17, 42). Together, these results suggest that the HP1 family of proteins may serve as one component of the KRAB–KAP-1 repression complex in vivo.

## DISCUSSION

The KRAB-ZFP family of zinc finger-containing transcriptional repressors has the potential for becoming the largest single class of human transcription factors which contain a common set of DNA binding ($C_2H_2$ zinc fingers) and effector (KRAB domain) motifs. It has been estimated that the human genome contains approximately 700 genes that encode $C_2H_2$ zinc finger proteins and that one-third or more of these contain KRAB domains (28). Their prevalence and conservation indicate the need for a molecular understanding of their mechanism of action. A significant advance was the discovery, purification, and cloning of a novel gene encoding a potential universal binding protein which functions as a corepressor for the KRAB domain, KAP-1. The present work represents the beginning of studies aimed at identifying downstream effector components operative in the KRAB–KAP-1 repression pathway.

**KAP-1 recruits HP-like proteins which mediate KRAB-ZFP repression.** This report demonstrates that KAP-1 interacts directly with members of the chromodomain- or chromo shadow domain-containing heterochromatin protein family including M31, M32, hHP1α, and hHP1γ. Efficient binding can be detected in solution-based chromatography assays and in EMSAs using recombinant GST-HP1 family proteins. The interaction is a direct protein-protein interaction between KAP-1 and HP1, as shown using highly purified recombinant versions of each. We have localized an HP1BD in KAP-1 and have shown that amino acid substitutions in the HP1BD abolish HP1 binding. The complex formed in vitro is stable to gel filtration

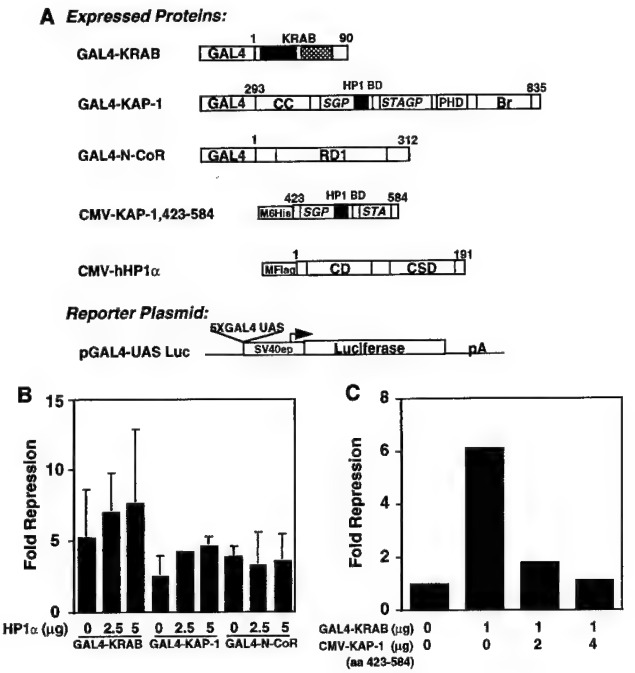



FIG. 10. Modulation of KRAB/KAP-1 repression activity by exogenous hHP1α and a dominant negative KAP-1. (A) Plasmids used for transfection. The indicated GAL4-DNA binding domain fusion proteins were expressed from the SV40 enhancer/promoter (SV40ep) in mammalian cells. The full-length hHP1α contained a FLAG epitope tag at the amino terminus. The fragment of KAP-1 (aa 423 to 584) which contains the HP1BD was expressed as an $NH_2$-terminal histidine fusion protein from a CMV vector. The SV40 promoter-based luciferase (Luc) reporter plasmid contains five synthetic GAL4-UAS sites and was used in all transfections. CC, coiled coil; Br, bromodomain; RD1, repression domain 1; CD, chromodomain; CSD, chromo shadow domain; pA, poly(A) site. (B) One microgram of each GAL4 expression plasmid was cotransfected into NIH 3T3 cells with increasing amounts of the hHP1α plasmid. Fold repression was calculated by comparing normalized luciferase activities with that of cells transfected in the absence of a GAL4 effector protein. The error bars represent the standard errors observed for three independent transfections, each performed in duplicate. (C) A GAL4-KRAB expression plasmid was cotransfected with increasing amounts of CMV–KAP-1 (aa 423 to 584), which encodes the putative HP1BD. Note that GAL4-KRAB repression activity is almost completely abolished in the presence of 4 μg of the CMV–KAP-1 (aa 423 to 584) plasmid, suggesting that the HP1BD is titrating endogenous HP1 proteins, thus relieving GAL4-KRAB-mediated repression via a squelching mechanism.

chromatography and alters the apparent native molecular mass of a KAP-1-containing complex from 560 kDa to around 800 kDa (Fig. 7). There are several explanations for this apparent increase: (i) multiple molecules of M32 interact with KAP-1 and/or each other, (ii) additional heterochromatin proteins from the nuclear extract bind M32 bound to the KAP-1 complex, and (iii) both heterochromatin proteins and other interacting proteins interact with the new KAP-1–GST-M32 complex to shift the endogenous complex. Given the demonstration that chromodomains of heterochromatin proteins are able to homo and hetero-oligomerize (10, 41, 42), we favor the possibility that between two and four M32 molecules interact with a KAP-1 complex. We have also succeeded in using EMSA to reconstitute a stable, quaternary complex containing DNA, a KRAB domain, KAP-1, and heterochromatin proteins. This finding suggests that chromodomain proteins, exemplified by M32, do not displace KAP-1 from the KRAB domain upon binding. Furthermore, KAP-1–M32 interaction does not abrogate DNA binding by KRAB domain-containing DNA binding proteins. Together, these observations suggest that endogenous KAP-1 has the capability of recruiting HP1 family members to a gene bound by a KRAB-ZFP protein. Consistent with this, we observed that transient expression of hHP1α augmented KRAB–KAP-1-mediated repression. Furthermore, overexpression of a segment of KAP-1 which contains the HP1BD abolished KRAB-mediated repression. Finally, mutations in the HP1BD which abolish HP1 binding significantly reduce the repression potential of a GAL4–KAP-1 protein. Thus, recruitment of heterochromatin proteins may play an important role in KRAB–ZFP–KAP-1-dependent repression (Fig. 12). Consistent with this hypothesis is the observation that HP1 family members themselves are potent, DNA binding-dependent repressors when tethered to a template via fusion to heterologous DNA binding domains (30).

**Heterochromatinization may be only part of KAP-1-mediated repression.** The predicted region of KAP-1 required for binding to HP1 proteins lies between the coiled-coil domain and the PHD (Fig. 1). Consistent with this, analysis of this central domain of KAP-1 revealed a region of amino acid se-

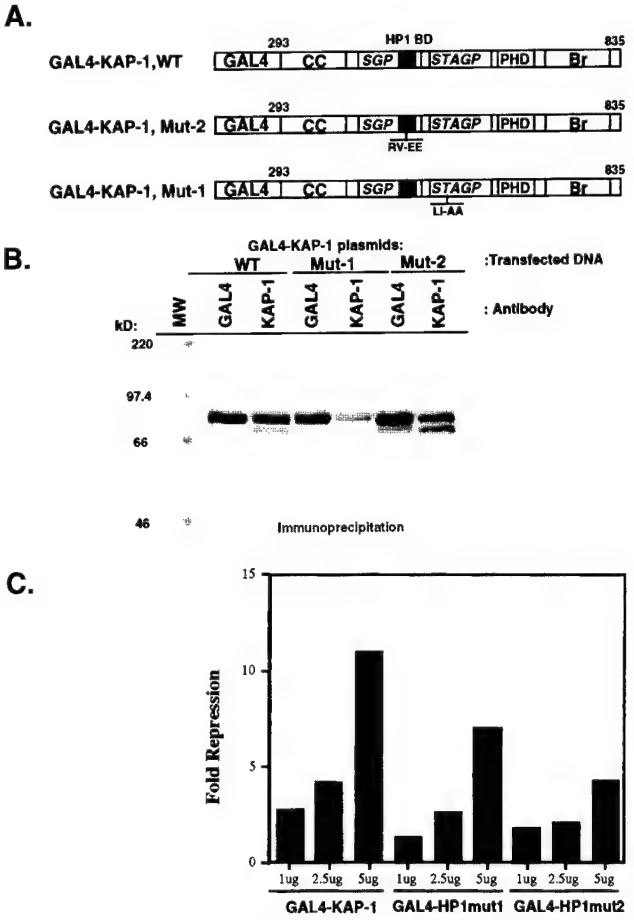



FIG. 11. Mutations in the HP1BD reduce the intrinsic repression activity of the KAP-1 protein. (A) The plasmids used for transfection include GAL4 fusions to the wild-type (WT) KAP-1 or the Mut1 and Mut2 versions depicted in Fig. 5. Each plasmid (for notation, see the legend to Fig. 10) was tested for stable expression via transfection into COS-1 cells followed by immunoprecipitation analyses using both anti-KAP-1 and anti-GAL4 sera (B). These plasmids were then used in transfection assays with the 5XGAL4-UAS-SV40 luciferase reporter plasmid (Fig. 10A). (C) The wild-type (WT) GAL4–KAP-1 protein displayed potent repression of the reporter plasmid. However, the heterologous Mut2 protein, which does not bind HP1 proteins (Fig. 5) is reduced in repression activity, while the heterologous Mut1 protein displayed an intermediate effect on KAP-1-mediated repression. MW, molecular weight markers.

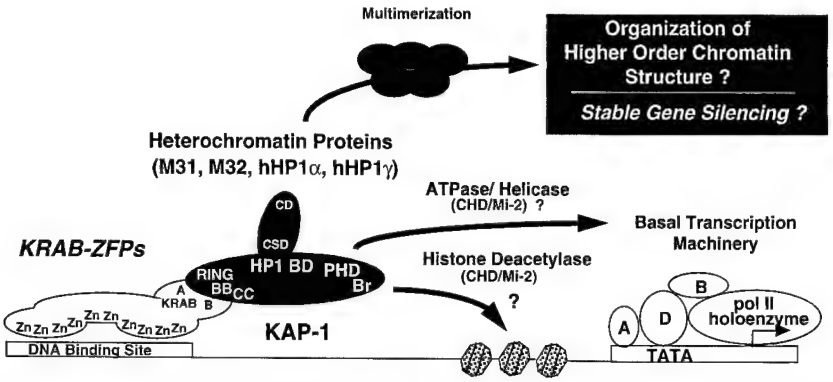



FIG. 12. Summary of protein-protein interactions identified in the KRAB-ZFP–KAP-1 repression pathway. We propose that the KRAB-ZFP family of transcriptional repressors function in part as sequence-specific DNA binding proteins which recruit the KAP-1 corepressor to a target gene. This interaction is dependent on the RBCC domain of KAP-1. Together, the HP1BD, PHD, and bromodomain comprise the surfaces which mediate gene silencing via interaction with the indicated potential partners. This work has defined members of the HP1 family of heterochromatin proteins as likely downstream effectors of KAP-1 which may mediate the assembly of stable, higher-order silenced domains in the eukaryotic nucleus. For notation, see the legend to Fig. 10. pol, polymerase.

quence homology termed the HP1BD, shared with the mTIF1α coactivator protein, which was previously shown to bind the mHP1α and M31 heterochromatin proteins (30). This finding suggests that the small-signature amino acid sequence motif may be responsible for protein-protein interactions involving HP1 family members and may be useful in identifying other components which are direct targets bound by HP1 family members. While the HP1BD is sufficient for binding of HP1 family members, we have previously shown that significant repression activity is also exhibited by a GAL4 fusion which contains only the PHD and bromodomains (42). One interpretation of these data is that multiple surfaces of KAP-1 contribute to the protein-protein interactions required for repression and that multiple mechanisms are involved. Paradigms for this scenario have been found for other large nuclear corepressors such as N-CoR, SMRT, Sin3, Mad, and YY1 (see reference 26 and references therein). These corepressor proteins appear to serve as molecular scaffolds for nucleation of multicomponent complexes required for transcriptional repression, the primary specificities of which are provided by the DNA binding components of the complexes. Molecular mechanisms for the actions of these complexes have come from the study of murine Sin3-HDAC complexes containing histone deacetylase activity (26). These complexes apparently mediate localized chromatin changes which occur at the level of the nucleosome. As a consequence of deacetylating core histone, there is a resulting increase in repression of transcription. However, the finding that KAP-1 binds to heterochromatin proteins implies that the KRAB–KAP-1 repression system may function at a different level of chromatin organization than simple histone modification.

**Potential for local KRAB–KAP-1/HP1 euchromatin heterochromatic regions.** Using MAbs to M31 and M32, we have shown that these proteins colocalize with KAP-1 within interphase nuclei (Fig. 8 and 9). These data lend support to the suggestion that the KRAB-ZFP–KAP-1 complex and HP1-like proteins are likely to interact in vivo. A thorough inspection of the cells immunostained with specific antibodies against KAP-1 revealed a dynamic staining pattern which may correlate with the cell cycle. Particularly striking is the localization of KAP-1 to distinct territories within the nuclei of some cells. Moreover, some of these territories overlap with the localization patterns of M31 and M32. This colocalization within subnuclear territories between KAP-1 and either M31 or M32 is reminiscent of the observations found with Ikaros, a transcriptional regulator that is essential for lymphoid development. In cells where the lymphoid cell-specific gene CD4 is inactive, the gene is localized, with Ikaros, around centromeric heterochromatin (8). By contrast, the active gene is localized elsewhere in the nucleoplasm. As for Ikaros, KRAB-ZFP binding to a gene may recruit the gene to a heterochromatic territory or form a new, localized heterochromatic domain around the gene, both of which would lead to its stable silencing.

In summary, the KRAB-ZFP family is a large set of proteins, each of which contains multiple contiguous $C_2H_2$ zinc fingers and associated KRAB repression domains (Fig. 12). These arrays of zinc fingers may mediate sequence-specific DNA binding, although recent evidence suggests that not all fingers in these long-array ZFPs participate in direct DNA contact. The KRAB domain recruits the KAP-1 corepressor to the DNA template bound by a KRAB-ZFP, binding directly to the RBCC domain of KAP-1. Repression of transcription by a KRAB-ZFP requires recruitment of KAP-1. KAP-1 homo-oligomerizes (unpublished data) upon binding to the KRAB domain and exhibits intrinsic repression activity through the COOH-terminal domains that include the HP1BD, PHD, and bromodomain. The HP1BD of KAP-1 binds to heterochromatic and euchromatic HP1 proteins, resulting in either recruitment of the KRAB-ZFP-bound target gene to heterochromatic chromosomal territories in the nucleus or formation of local heterochromatic chromosomal regions and subsequent silencing of gene expression. Homo- or hetero-oligomerization of HP1 protein family members may play a role in transient and/or stable gene repression, and it is interesting to speculate that this may be catalyzed at a locus by the KRAB-ZFP–KAP-1 system. Additional potential components of the system which still must be considered include the chromodomain-containing CHD/Mi-2 protein, an ATPase/helicase which has been shown to both bind KAP-1 (unpublished data) and be an integral component of a histone deacetylase complex (52, 54, 58).

We suggest that the KAP-1-directed recruitment of HP1-like proteins may provide an example of a mechanism by which euchromatic genes may be locally silenced, through the formation of a heterochromatin-like complex (19, 23). Thus, KAP-1 binding of a KRAB domain-containing protein provides the specificity for the assembly of a mammalian HP1-containing heterochromatin-like complex to a distinct site within the genome, as previously suggested (23), and may be mechanistically related to the mechanisms whereby Hunchback and dYY1 proteins (pleiohomeotic [7]) direct assembly of the Pc-

Group complexes in *Drosophila* (19, 38, 45). Moreover, the formation of a heterochromatin-like complex by the KRAB-ZFP–KAP-1 complex may lead to the dynamic recruitment of silenced genes to a repressive chromosomal environment. The driving energy for such recruitment may come from the complementarity shared between the chromatin components of the repressor complexes, a feature recently described in a model proposed to explain such associations of repressed chromosomal domains (46).

## ACKNOWLEDGMENTS

R.F.R., D.C.S., and K.A. contributed equally to the article and should be considered joint first authors.

We thank David E. Jensen for preparation and purification of the 6HisGAL4-KRAB protein, and we thank Qinwu Liu and Gerd Maul, Wistar Institute Microscopy Core Facility, for help in generating and deciphering the immunofluorescence images. We thank M. Lazar for the GAL4–N-CoR plasmid and many helpful discussions. We thank H. Worman for the hHP1γ plasmid.

R.F.R., W.J.F., and D.C.S. were supported by Wistar Basic Cancer Research training grant CA 09171. F.J.R. was supported in part by National Institutes of Health grant CA 52009, Core grant CA 10815, Core grant DK50306, and grants DK 49210, GM 54220, DAMD 17-96-1-6141, and ACS NP-954, the Irving A. Hansen Memorial Foundation, the Mary A. Rumsey Memorial Foundation, and the Pew Scholars Program in the Biomedical Sciences. P.B.S. was supported by the BBSRC.

## REFERENCES

1. **Aasland, R., and A. F. Stewart.** 1995. The chromo shadow domain, a second chromo domain in heterochromatin-binding protein 1, HP1. Nucleic Acids Res. **23:**3163–3173.
2. **Albagli, O., P. Dhordain, C. Deweindt, G. Lecocq, and D. Leprince.** 1995. The BTB/POZ domain: a new protein-protein interaction motif common to DNA- and actin-binding proteins. Cell Growth Differ. **6:**1193–1198.
3. **Auble, D. T., K. E. Hansen, C. G. Mueller, W. S. Lane, J. Thorner, and S. Hahn.** 1994. Mot1, a global repressor of RNA polymerase II transcription, inhibits TBP binding to DNA by an ATP-dependent mechanism. Genes Dev. **8:**1920–1934.
4. **Baniahmad, A., X. Leng, T. P. Burris, S. Y. Tsai, M. J. Tsai, and B. W. O'Malley.** 1995. The tau 4 activation domain of the thyroid hormone receptor is required for release of a putative corepressor(s) necessary for transcriptional silencing. Mol. Cell. Biol. **15:**76–86.
5. **Bellefroid, E. J., D. A. Poncelet, P. J. Lecocq, O. Revelant, and J. A. Martial.** 1991. The evolutionarily conserved Kruppel-associated box domain defines a subfamily of eukaryotic multifingered proteins. Proc. Natl. Acad. Sci. USA **88:**3608–3612.
6. **Britten, R. J., and E. H. Davidson.** 1969. Gene regulation for higher cells: a theory. Science **165:**349–357.
7. **Brown, J. L., D. Mucci, M. Whitely, M. L. Dirksen, and J. A. Kassis.** 1998. The Drosophila polycomb group gene pleiohomeotic encodes a DNA binding protein with homology to the transcription factor YY1. Mol. Cell **1:**1057–1064.
8. **Brown, K. E., S. S. Guest, S. T. Smale, K. Hahm, M. Merkenschlager, and A. G. Fisher.** 1997. Association of transcriptionally silent genes with Ikaros complexes at centromeric heterochromatin. Cell **91:**845–854.
9. **Brown, S. W.** 1966. Heterochromatin. Science **151:**417–425.
10. **Cowell, I. G., and C. A. Austin.** 1997. Self-association of chromo domain peptides. Biochim. Biophys. Acta **1337:**198–206.
11. **Dawson, S. R., D. L. Turner, H. Weintraub, and S. M. Parkhurst.** 1995. Specificity for the hairy/enhancer of split basic helix-loop-helix (bHLH) proteins maps outside the bHLH domain and suggests two separable modes of transcriptional repression. Mol. Cell. Biol. **15:**6923–6931.
12. **Eissenberg, J. C., T. C. James, D. M. Foster-Hartnett, T. Hartnett, V. Ngan, and S. C. Elgin.** 1990. Mutation in a heterochromatin-specific chromosomal protein is associated with suppression of position-effect variegation in Drosophila melanogaster. Proc. Natl. Acad. Sci. USA **87:**9923–9927.
13. **Eissenberg, J. C., G. D. Morris, G. Reuter, and T. Hartnett.** 1992. The heterochromatin-associated protein HP-1 is an essential protein in Drosophila with dosage-dependent effects on position-effect variegation. Genetics **131:**345–352.
14. **Fondell, J. D., F. Brunel, K. Hisatake, and R. G. Roeder.** 1996. Unliganded thyroid hormone receptor alpha can target TATA-binding protein for transcriptional repression. Mol. Cell. Biol. **16:**281–287.
15. **Fondell, J. D., A. L. Roy, and R. G. Roeder.** 1993. Unliganded thyroid hormone receptor inhibits formation of a functional preinitiation complex: implications for active repression. Genes Dev. **7:**1400–1410.
16. **Fredericks, W. J., N. Galili, S. Mukhopadhyay, G. Rovera, J. Bennicelli, F. G. Barr, and F. J. Rauscher III.** 1995. The PAX3-FKHR fusion protein created by the t(2;13) translocation in alveolar rhabdomyosarcomas is a more potent transcriptional activator than PAX3. Mol. Cell. Biol. **15:**1522–1535.
17. **Friedman, J. R., W. J. Fredericks, D. E. Jensen, D. W. Speicher, X. P. Huang, E. G. Neilson, and F. J. Rauscher III.** 1996. KAP-1, a novel corepressor for the highly conserved KRAB repression domain. Genes Dev. **10:**2067–2078.
18. **Furuta, K., E. K. L. Chan, K. Kiyosawa, G. Reimer, C. Luderschmidt, and E. M. Tan.** 1997. Heterochromatin protein HP1Hsbeta (p25beta) and its localization with centromeres in mitosis. Chromosoma **106:**11–19.
19. **Gaunt, S. J., and P. B. Singh.** 1990. Homeogene expression patterns and chromosomal imprinting. Trends Genet. **6:**208–212.
20. **Grimes, H. L., C. B. Gilks, T. O. Chan, S. Porter, and P. N. Tsichlis.** 1996. The Gfi-1 protooncoprotein represses Bax expression and inhibits T-cell death. Proc. Natl. Acad. Sci. USA **93:**14569–14573.
21. **Guarente, L.** 1995. Transcriptional coactivators in yeast and beyond. Trends Biochem. Sci. **20:**517–521.
22. **Hamvas, R. M., W. Reik, S. J. Gaunt, S. D. Brown, and P. B. Singh.** 1992. Mapping of a mouse homolog of a heterochromatin protein gene to the X chromosome. Mamm. Genome **2:**72–75.
23. **Horsley, D., A. Hutchings, G. W. Butcher, and P. B. Singh.** 1996. M32, a murine homologue of Drosophila heterochromatin protein 1 (HP1), localises to euchromatin within interphase nuclei and is largely excluded from constitutive heterochromatin. Cytogenet. Cell Genet. **73:**308–311.
24. **Huang, X.-P., and F. J. Rauscher III.** Unpublished data.
25. **James, T. C., and S. C. Elgin.** 1986. Identification of a nonhistone chromosomal protein associated with heterochromatin in *Drosophila melanogaster* and its gene. Mol. Cell. Biol. **6:**3862–3872.
26. **Kadosh, D., and K. Struhl.** 1998. Histone deacetylase activity of Rpd3 is important for transcriptional repression in vivo. Genes Dev. **12:**797–805.
27. **Kim, S. S., Y. M. Chen, E. O'Leary, R. Witzgall, M. Vidal, and J. V. Bonventre.** 1996. A novel member of the RING finger family, KRIP-1, associates with the KRAB-A transcriptional repressor domain of zinc finger proteins. Proc. Natl. Acad. Sci. USA **93:**15299–15304.
28. **Klug, A., and J. W. Schwabe.** 1995. Protein motifs 5. Zinc fingers. FASEB J. **9:**597–604.
29. **Lassar, A. B., R. L. Davis, W. E. Wright, T. Kadesch, C. Murre, A. Voronova, D. Baltimore, and H. Weintraub.** 1991. Functional activity of myogenic HLH proteins requires hetero-oligomerization with E12/E47-like proteins in vivo. Cell **66:**305–315.
30. **Le Douarin, B., A. L. Nielsen, J. M. Garnier, H. Ichinose, F. Jeanmougin, R. Losson, and P. Chambon.** 1996. A possible involvement of TIF1 alpha and TIF1 beta in the epigenetic control of transcription by nuclear receptors. EMBO J. **15:**6701–6715.
31. **Le Douarin, B., A. L. Nielsen, J. You, P. Chambon, and R. Losson.** 1997. TIF1 alpha: a chromatin-specific mediator for the ligand-dependent activation function AF-2 of nuclear receptors? Biochem. Soc. Trans. **25:**605–612.
32. **Le Douarin, B., E. vom Baur, C. Zechel, D. Heery, M. Heine, V. Vivat, H. Gronemeyer, R. Losson, and P. Chambon.** 1996. Ligand-dependent interaction of nuclear receptors with potential transcriptional intermediary factors (mediators). Philos. Trans. R. Soc. Lond. Ser. B **351:**569–578.
33. **Le Douarin, B., C. Zechel, J. M. Garnier, Y. Lutz, L. Tora, B. Pierrat, D. Heery, H. Gronemeyer, P. Chambon, and R. Losson.** 1995. The N-terminal part of TIF1, a putative mediator of ligand-dependent activation function (AF-2) of nuclear receptors, is fused to B-raf in the oncogenic protein T18. EMBO J. **14:**2020–2033.
34. **Margolin, J. F., J. R. Friedman, W. K. Meyer, H. Vissing, H. J. Thiesen, and F. J. Rauscher III.** 1994. Kruppel-associated boxes are potent transcriptional repression domains. Proc. Natl. Acad. Sci. USA **91:**4509–4513.

34a. **Maul, G. G.** Unpublished data.

35. **Maul, G. G., D. E. Jensen, A. M. Ishov, M. Herlyn, and F. J. Rauscher III.** 1998. Nuclear redistribution of BRCA1 during viral infection. Cell Growth Differ. **9:**743–755.
36. **Moosmann, P., O. Georgiev, B. Le Douarin, J. P. Bourquin, and W. Schaffner.** 1996. Transcriptional repression by RING finger protein TIF1 beta that interacts with the KRAB repressor domain of KOX1. Nucleic Acids Res. **24:**4859–4867.
37. **Nicol, L., and P. Jeppesen.** 1994. Human autoimmune sera recognize a conserved 26 kD protein associated with mammalian heterochromatin that is homologous to heterochromatin protein 1 of Drosophila. Chromosome Res. **2:**245–253.
38. **Paro, R.** 1990. Imprinting a determined state into the chromatin of Drosophila. Trends Genet. **6:**416–421.
39. **Paro, R., and D. S. Hogness.** 1991. The Polycomb protein shares a homologous domain with a heterochromatin-associated protein of Drosophila. Proc. Natl. Acad. Sci. USA **88:**263–267.
40. **Pearce, J. J., P. B. Singh, and S. J. Gaunt.** 1992. The mouse has a Polycomb-like chromobox gene. Development **114:**921–929.
41. **Platero, J. S., T. Hartnett, and J. C. Eissenberg.** 1995. Functional analysis of the chromo domain of HP1. EMBO J. **14:**3977–3986.

42. **Rauscher, F. J., III, J. R. Friedman, R. F. Ryan, D. C. Schultz, and W. J. Fredericks.** Unpublished data.
43. **Sauer, F., J. D. Fondell, Y. Ohkuma, R. G. Roeder, and H. Jäckle.** 1995. Control of transcription by Kruppel through interactions with TFIIB and TFIIE beta. Nature **375:**162–164.
44. **Saunders, W. S., C. Chue, M. Goebl, C. Craig, R. F. Clark, J. A. Powers, J. C. Eissenberg, S. C. Elgin, N. F. Rothfield, and W. C. Earnshaw.** 1993. Molecular cloning of a human homologue of Drosophila heterochromatin protein HP1 using anti-centromere autoantibodies with anti-chromo specificity. J. Cell Sci. **104:**573–582.
45. **Singh, P. B.** 1994. Molecular mechanisms of cellular determination: their relation to chromatin structure and parental imprinting. J. Cell Sci. **107:** 2653–2668.
46. **Singh, P. B., and N. S. Huskisson.** 1998. Chromatin complexes as aperiodic microcrystalline arrays that regulate genome organisation and expression. Dev. Genet. **22:**85–99.
47. **Singh, P. B., J. R. Miller, J. Pearce, R. Kothary, R. D. Burton, R. Paro, T. C. James, and S. J. Gaunt.** 1991. A sequence motif found in a Drosophila heterochromatin protein is conserved in animals and plants. Nucleic Acids Res. **19:**789–794.
48. **Spofford, J. B.** 1976. Position-effect variegation in *Drosophila*, p. 955–1018. *In* M. Ashburner and E. Novitski (ed.), The genetics and biology of Drosophila, vol. 1c. Academic Press, New York, N.Y.
49. **Sugimoto, K., T. Yamada, Y. Muro, and M. Himeno.** 1996. Human homolog of *Drosophila* heterochromatin-associated protein 1 (HP1) is a DNA-binding protein which possesses a DNA-binding motif with weak similarity to that of human centromere protein C (CENP-C). J. Biochem. **120:**153–159.
50. **Tartof, K. D., and M. Bremer.** 1990. Mechanisms for the construction and developmental control of heterochromatin formation and imprinted chromosome domains. Dev. Suppl., p. 35–45.
51. **Taunton, J., C. A. Hassig, and S. L. Schreiber.** 1996. A mammalian histone deacetylase related to the yeast transcriptional regulator Rpd3p. Science **272:**408–411.
52. **Tong, J. K., C. A. Hassig, G. R. Schnitzler, R. E. Kingston, and S. L. Schreiber.** 1998. Chromatin deacetylation by an ATP-dependent nucleosome remodelling complex. Nature **395:**917–921.
53. **Um, M., C. Li, and J. L. Manley.** 1995. The transcriptional repressor even-skipped interacts directly with TATA-binding protein. Mol. Cell. Biol. **15:** 5007–5016.
54. **Wade, P. A., P. L. Jones, D. Vermaak, and A. P. Wolffe.** 1998. A multiple subunit Mi-2 histone deacetylase from Xenopus laevis cofractionates with an associated Snf2 superfamily ATPase. Curr. Biol. **8:**843–846.
55. **Wolffe, A. P.** 1996. Histone deacetylase: a regulator of transcription. Science **272:**371–372.
56. **Wreggett, K. A., F. Hill, P. S. James, A. Hutchings, G. W. Butcher, and P. B. Singh.** 1994. A mammalian homologue of Drosophila heterochromatin protein 1 (HP1) is a component of constitutive heterochromatin. Cytogenet. Cell Genet. **66:**99–103.
57. **Ye, Q., and H. J. Worman.** 1996. Interaction between an integral protein of the nuclear envelope inner membrane and human chromodomain proteins homologous to Drosophila HP1. J. Biol. Chem. **271:**14653–14656.
58. **Zhang, Y., G. LeRoy, H.-P. Seelig, W. S. Lane, and D. Reinberg.** 1998. The dermatomyositis-specific autoantigen Mi2 is a component of a complex containing histone deacetylase and nucleosome remodeling activities. Cell **95:** 279–289.

Article No. jmbi.1999.3402 available online at http://www.idealibrary.com on IDEAL®  J. Mol. Biol. (2000) 00, 1–24

JMB

# Reconstitution of the KRAB-KAP-1 Repressor Complex: A Model System for Defining the Molecular Anatomy of RING-B Box-coiled-coil Domain-mediated Protein-protein Interactions

**Hongzhuang Peng, Gillian E. Begg, David C. Schultz, Josh R. Friedman David E. Jensen, David W. Speicher and Frank J. Rauscher III**

*The Wistar Institute, 3601 Spruce Street, Philadelphia, PA 19104, USA*

**Corresponding author*

The KRAB domain is a 75 amino acid residue transcriptional repression module commonly found in eukaryotic zinc-finger proteins. KRAB-mediated gene silencing requires binding to the corepressor KAP-1. The KRAB:KAP-1 interaction requires the RING-B box-coiled coil (RBCC) domain of KAP-1, which is a widely distributed motif, hypothesized to be a protein-protein interface. Little is known about RBCC-mediated ligand binding and the role of the individual sub-domains in recognition and specificity. We have addressed these issues by reconstituting and characterizing the KRAB:KAP-1-RBCC interaction using purified components. Our results show that KRAB binding to KAP-1 is direct and specific, as the related RBCC domains from TIF1α and MID1 do not bind the KRAB domain. A combination of gel filtration, analytical ultracentrifugation, chemical cross-linking, non-denaturing gel electrophoresis, and site-directed mutagenesis techniques has revealed that the KAP-1-RBCC must oligomerize likely as a homo-trimer in order to bind the KRAB domain. The RING finger, B2 box, and coiled-coil region are required for oligomerization of KAP-1-RBCC and KRAB binding, as mutations in these domains concomitantly abolished these functions. KRAB domain binding stabilized the homo-oligomeric state of the KAP-1-RBCC as detected by chemical cross-linking and velocity sedimentation studies. Mutant KAP-1-RBCC molecules hetero-oligomerize with the wild-type KAP-1, but these complexes were inactive for KRAB binding, suggesting a potential dominant negative activity. Substitution of the coiled-coil region with heterologous dimerization, trimerization, or tetramerization domains failed to recapitulate KRAB domain binding. Chimeric KAP-1-RBCC proteins containing either the RING, RING-B box, or coiled-coil regions from MID1 also failed to bind the KRAB domain. The KAP-1-RBCC mediates a highly specific, direct interaction with the KRAB domain, and it appears to function as an integrated, possibly cooperative structural unit wherein each sub-domain contributes to oligomerization and/or ligand recognition. These observations provide the first principles for RBCC domain-mediated protein-protein interaction and have implications for identifying new ligands for RBCC domain proteins.



*Keywords:* transcriptional repression; KRAB domain; RBCC domain; RING finger; protein-protein interaction

Present addresses: G. E. Begg, Victor Chang Cardiac Research Institute, Darlinghurst, NSW 2010, Australia; D. E. Jensen, GlaxoWellcome, Inc. Research Triangle Park, North Carolina, 27709, USA.

Abbreviations used: KRAB, Krüppel associated box; RBCC, RING finger, B boxes, and coiled-coil region; Ni-NTA, $Ni^{2+}$-nitrilo-tri-acetic acid; ND-PAGE, Non-denaturing polyacrylamide gel electrophoresis; EMSA, electrophoretic mobility shift assay; EGS, ethylene glycol bis (succinimidylsuccinate); bv, baculovirus; HSF, heat shock transcription factor; ZFPs, zinc-finger proteins.

E-mail address for the corresponding author: rauscher@wistar.upenn.edu

0022-2836/00/0000001–24 $30.00/0

## Introduction

The recurrent theme that has emerged from the structural study of proteins involved in signal transduction and transcriptional regulation is that multiple independently folded globular domains in these proteins often cooperate in macromolecular recognition. These domains are often recognizable by conserved signature amino acid sequence motifs and their organization within a protein often is the first clue about protein function. The bHLH and bZIP DNA-binding proteins have provided an important paradigm for protein multimerization and macro-molecular recognition: each protein encodes adjacent, functionally separable dimerization and DNA-binding regions (Baxevanis & Vinson, 1993; Hurst, 1995). The dimerization function, which is mediated by a coiled-coil motif, is required to correctly orient the basic regions for interaction with DNA. With these principles in mind, we have been characterizing the RING-B box-coiled-coil (RBCC) family of nuclear regulatory proteins, which have been predicted to contain functionally separable multimerization and protein recognition domains (Borden, 1998).

The RBCC family of proteins contains the tripartite motif that most likely functions as a cooperative protein-protein interaction motif (Borden, 1998) (Figure 1(a) and (b)). The definitive element of the tripartite motif is the RING finger, which is a cysteine-rich motif of the form C3HC4 and binds two molecules of zinc in a unique cross-braced ligation system (Barlow *et al.*, 1994; Bellon *et al.*, 1997; Borden *et al.*, 1995a). The RING finger is found almost exclusively in the $NH_2$-terminal position in RBCC proteins and is likely to contribute either specificity and/or multimerization properties to the tripartite motif. More than 100 proteins encode RING finger, and mutational analyses have confirmed the requirement for the RING finger for proper biological function (for reviews, see Borden, 1998; Saurin *et al.*, 1996; Boddy *et al.*, 1997; Cao *et al.*, 1997). RING finger structures suggest that a common hydrophobic core is formed as a result of zinc-chelation and that the variable sequence between the conserved ligation residues provides specificity for protein recognition.

The second signature motif of the RBCC domain is the B-box, which is also a cysteine-rich zinc-binding motif of the form CHC3H2 (Reddy & Etkin, 1991) (Figure 1(b)). The B-box is an independently folded globular domain. Two B-box motifs are often found immediately carboxy-terminal to the RING finger in the RBCC domain. The NMR structure of the XNF7 B-box showed that only one zinc atom was bound with the other potential chelation residues unoccupied (Borden *et al.*, 1995b). XNF7 is required for embryonic development *via* a chromatin-binding function that requires the integrity of the B-box (Bellini *et al.*, 1995).

The third RBCC signature motif is a coiled-coil domain that is comprised of appropriately spaced hydrophobic residues predicted to form an extended α-helical region (for a review, see Lupas, 1996) (Figure 1(b)). Multiple individual regions within the predicted coiled-coil region display amphipathic character. Moreover, many of the biological properties of RBCC proteins have been shown to be dependent on multimerization *via* this coiled-coil region (Saurin *et al.*, 1996).

Thus, RBCC proteins are defined by this tripartite motif, which is found in a spatially restricted organization in each protein. Clues to the function and/or potential ligands for RBCC proteins have come mainly from the studies of their biological properties. Most notably, RBCC proteins have been defined as potential proto-oncogenes: RFP, TIF1α, and PML (Goddard *et al.*, 1991; Kakizuka *et al.*, 1991; Le Douarin *et al.*, 1995; Miki *et al.*, 1991; Takahashi *et al.*, 1985). The RBCC domain of each of these proteins is fused to heterologous proteins (as a result of chromosomal translocation) and results in the creation of oncogenes. The RBCC domain of PML protein is fused to the retinoic acid receptor alpha (RARα) as a result of the t(15:17) translocation (Borden *et al.*, 1996; Goddard *et al.*, 1991; Kakizuka *et al.*, 1991), resulting in aberrant regulation of RAR target genes and a block in myelocytic differentiation (Grignani *et al.*, 1996). The PML-RARα fusion homodimerizes and heterodimerizes with the wild-type PML and RXR: both of these activities require the intact coiled-coil region of the RBCC domain (Dyck *et al.*, 1994; Perez *et al.*, 1993; Weis *et al.*, 1994). Multimerization appears to sequester wild-type RAR/RXR receptors into discrete subnuclear domains termed ND10 or PODs (for a review, see Maul, 1998; Sternsdorf *et al.*, 1997). The wild-type PML protein functions as a growth suppressor *via* induction of apoptosis (Borden *et al.*, 1997). This activity requires both localization to ND10/PODs and cooperation between the RING, B-boxes and coiled-coil motifs. Thus, in this biological context, the PML-RBCC domain functions as a cooperative unit that presumably mediates specific protein-protein interactions.

The Krüppel associated box (KRAB) domain was originally identified as a conserved motif at the $NH_2$-terminus of zinc-finger proteins (ZFPs) (Bellefroid *et al.*, 1991) and was then shown to be a potent, DNA-binding-dependent transcriptional repression module (Margolin *et al.*, 1994; Pengue *et al.*, 1995; Vissing *et al.*, 1995; Witzgall *et al.*, 1994). The KRAB domain homology consists of approximately 75 amino acid residues and is predicted to fold into two amphipathic helices (Figure 1(c)). The minimal repression module is approximately 45 amino acid residues, and substitutions for conserved residues abolish repression (Margolin *et al.*, 1994). More than ten independently encoded KRAB domains have been demonstrated to be potent repressors, suggesting that this activity is a common property. Our laboratory purified and cloned KAP-1 as a protein that binds to the KRAB repression domain (Friedman *et al.* 1996). Structurally, KAP-1 possesses all the

signature motifs for an emerging superfamily of transcriptional regulators, including TIF1α and TIF1γ (Friedman *et al.* 1996; Le Douarin *et al.*, 1995; Venturini *et al.*, 1999) (Figure 1(a)). Further studies indicated that KAP-1 functions as a universal corepressor for KRAB domain proteins (Friedman *et al.*, 1996). KAP-1 is a 97 kDa nuclear phosphoprotein that possesses a canonical RBCC domain in the $NH_2$ terminus, and appears to be both necessary and sufficient for the KRAB domain binding. The carboxy terminus of KAP-1 includes a PHD domain and a bromodomain, and this region functions as a repressor when tethered to DNA using a heterologous DNA-binding domain, and mutations in either domain weaken this repression activity (unpublished results).

It has been firmly established that KAP-1 is required for KRAB domain-mediated repression. The evidence includes: (1) KAP-1 binds to multiple KRAB repression domains both *in vitro* and *in vivo*; (2) KRAB domain mutations that abolish repression decrease or eliminate KAP-1 binding; (3) overexpression of KAP-1 enhances KRAB-mediated repression; (4) the KRAB domain does not repress in cells that lack KAP-1. These results support a model in which KRAB-ZFPs bind a gene and repress transcription of the gene by recruiting the KAP-1 corepressor *via* the RBCC domain. Thus, the KRAB-KAP-1 interaction serves as a specific example from which to dissect molecular determinants for RBCC-ligand interactions. We have reconstituted the KRAB-KAP-1 complex using purified components and have employed a comprehensive set of biochemical approaches to analyze this interaction. Our data indicate that KAP-1-RBCC oligomerization is required for KRAB domain binding and that all three components of the tripartite motif cooperate in KRAB recognition.

## Results

### Direct interaction between the KAP-1-RBCC domain and the KOX1-KRAB domain

Previous studies using $NH_2$-terminal truncated KAP-1 proteins have suggested that the RING finger, B boxes, and coiled-coil structure are necessary for binding to the KRAB domain (Friedman *et al.*, 1996). However, all of these studies were performed using indirect protein-protein interaction assays either in complex mixtures *in vitro*, or in yeast two-hybrid assays. We sought to reconstitute the KRAB-KAP-1 interaction using purified recombinant components. The proteins were overexpressed using bacterial or baculovirus systems (see below) and purified using Ni-NTA chromatography to near-homogeneity (Figure 2(a)). The KAP-1-RBCC proteins produced in *Escherichia coli* were soluble and were purified under non-denaturing conditions. SDS-PAGE analysis (Figure 2(a)) revealed that the KAP-1-RBCC protein migrates with an apparent molecular mass of 46 kDa, a value that is almost identical with its predicted molecular mass of 45.9 kDa. The KAP-1-RBCC mutants show similar degrees of purity and migrate identically in SDS-PAGE (Figure 2(a)).

The KRAB domain that we have utilized in these studies corresponds to amino acid residues 1-90 of the KOX1 zinc finger protein (Figure 1(c)). Although the KRAB domain is found in more than 100 independent proteins, we selected the KRAB domain from KOX1 for the following reasons: (1) the KOX1-KRAB domain was originally utilized to isolate the KAP-1 corepressor, and mutations in this domain that concomitantly abolish repression and KAP-1 binding are well characterized (Friedman *et al.*, 1996). (2) The KOX1-KRAB domain is highly expressed in *E. coli* autonomously or as a GAL4 fusion and is well behaved in protein reconstitution assays. (3) The GAL4-KOX1-KRAB 1-90 fusion (designated GAL4-KRAB) is a potent, KAP-1-dependent and DNA binding-dependent transcriptional repressor *in vivo* (Margolin *et al.*, 1994). The wild-type GAL4-KRAB and a mutant version (DV18,19AA, which lacks repression activity) were expressed and purified from *E. coli* (Figure 2(a)). To detect whether the purified KAP-1-RBCC protein was able to form a stable complex with a DNA-bound KRAB domain, we used the electrophoretic mobility shift-supershift assay (EMSA). Binding of the GAL4-KRAB protein to a canonical $^{32}$P-labeled synthetic oligonucleotide containing the GAL4 recognition sequence yielded an expected mobility shift (Figure 2(b), lane 1). When increasing amounts of purified KAP-1-RBCC protein were incubated with the GAL4-KRAB protein and DNA, a new mobility shift was observed (Figure 2(b), lanes 2-6). This supershift contains the ternary complex of DNA:GAL4-KRAB:KAP-1-RBCC, as characterized by antibody inhibition/supershift and competition with unlabeled competitor DNA (data not shown). Formation of this complex requires the GAL4-KRAB DNA binding subunit since no shift is seen with KAP-1-RBCC alone (data not shown). The GAL4-KRAB(DV-AA) protein, which shows severely impaired repression function *in vivo* readily bound to $^{32}$P-labeled GAL4 DNA but failed to generate the ternary complex in the presence of increasing amounts of KAP-1-RBCC protein (Figure 2(b), lanes 7-12), in agreement with previous reports (Friedman *et al.*, 1996). These data strongly suggest that the KAP-1-KRAB interaction is direct and highly specific, and that the RBCC region is necessary and sufficient for this interaction.

### The RING finger, B boxes, and the coiled-coil region are critical for the formation of the KRAB:KAP-1-RBCC complex

To define the sub-domains of the KAP-1-RBCC domain integral to its interaction with the KRAB domain, we made independent mutations in the RING finger, B boxes, and the coiled-coil region (Figure 1(d)). Mutant proteins were expressed and purified in *E. coli* (Figure 2(a)), and analyzed for

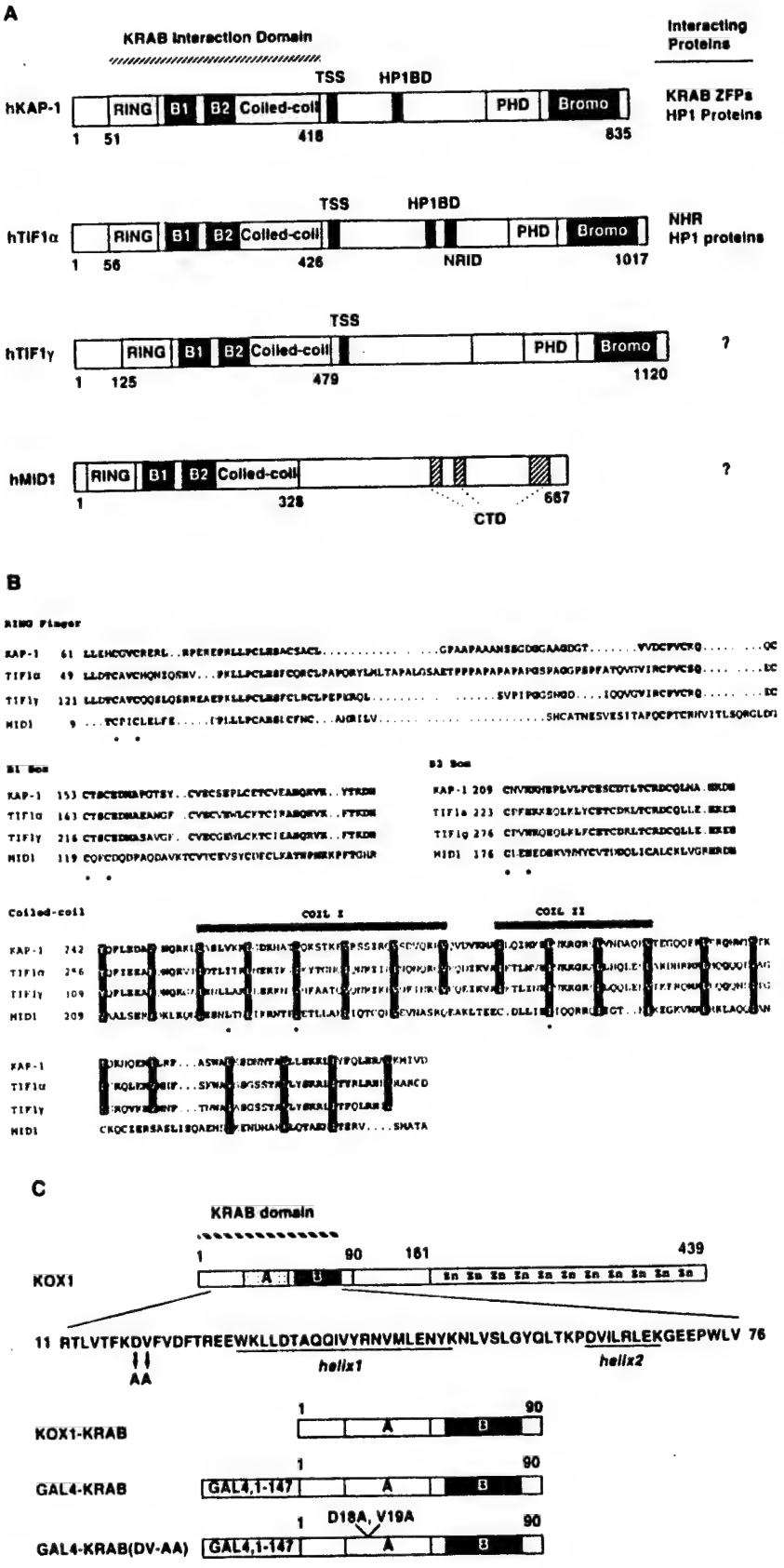



**Figure 1** (*legend opposite*)

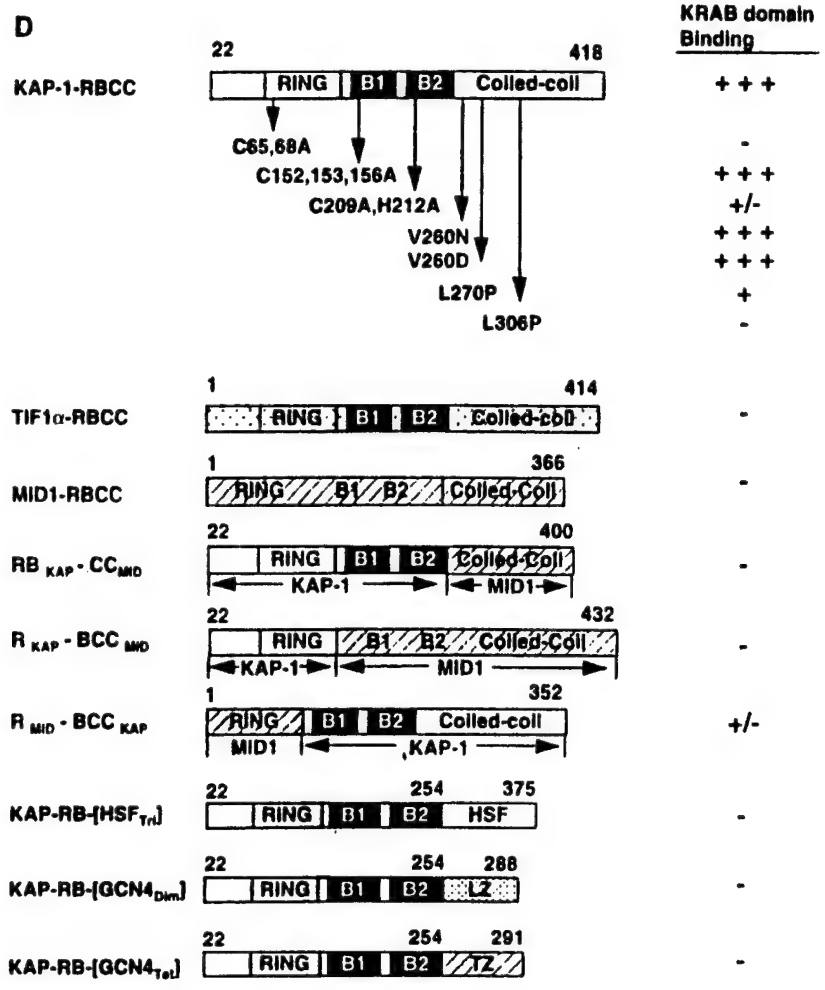



**Figure 1.** A diagram illustrating the architecture of the KAP-1/TIF1 family of transcriptional regulatory proteins and the KOX1/KRAB-ZFP proteins. (a) A representation of proteins containing the tripartite RBCC motif, including hKAP-1, hTIF1α, hTIF1γ, and hMID1. The numbers represent amino acid positions. RING, RING finger; B1 and B2, B-box; TSS, TIF1 signature sequence (Venturini *et al.*, 1999); HP1BD, HP1 binding domain (Ryan *et al.*, 1999); PHD, plant homeodomain; Bromo, bromo domain; NRID, nuclear receptor interaction domain; CTD, carboxy-terminal domain. The column to the right summarizes known protein-protein interactions: NHR, nuclear hormone receptor; HP1, heterochromatin protein-1. (b) Amino acid alignment of the RING finger, B boxes, and coiled-coil region of the hKAP-1, hTIF1α, hTIF1γ, and hMID1 proteins. The alignments were generated using the CLUSTAL program. The amino acid residues that are identical and/or conserved in all four proteins are in bold. The asterisks below indicate the amino acids targeted for mutational analysis in this study. The periods represent spaces introduced to obtain maximal alignment. The spacing between the B2-box and the first coil of KAP-1 is 15 amino acid residues. Within the coiled-coil domain, hydrophobic amino acid residues defining the heptad repeats of each putative coil are highlighted in black. The relative position of the first and the second predicted coils of the coiled-coil region are indicated by the gray bars above the sequence alignment. The numbers at the left refer to the amino acid positions in the corresponding proteins. Database accession numbers: hKAP-1 (998813); hTIF1α (AF119042); hTIF1γ (AF119043); hMID1 (Y13667). (c) A representation of the KOX1-ZFP and recombinant derivatives used in this study. The amino acid sequence of the minimal KRAB domain of KOX1 is illustrated. The position of the DV18,19AA mutation, which disrupts KRAB-mediated repression and association with KAP-1, is indicated (Margolin *et al.* 1994; Friedman *et al.* 1996). The KRAB domain of KOX1, residues 1-90, was expressed without a DNA-binding domain and as a heterologous fusion protein with the DNA-binding domain of GAL4, residues 1-147. (d) A diagram of the KAP-1-RBCC domain (wild-type and mutants) and organization of heterologous KAP-1-RBCC fusion proteins. The KAP-1-RBCC (CC65,68AA) is a double substitution in the RING finger. The KAP-1-RBCC (CCC152,153,156AAA) is a triple substitution in the B1 box. The KAP-1-RBCC (C209A,H212A) is a double substitution in the B2 box. The KAP-1-RBCC (V260N), KAP-1-RBCC (V260D), and KAP-1-RBCC (L270P) are single substitutions in the first predicted coil of the coiled-coil region. The KAP-1-RBCC (L306P) is a single substitution in the second predicted coil of the coiled-coil region. The TIF1α-RBCC domain consists of amino acid residues 1-414 of the TIF1α protein. The MID1-RBCC domain consists of residues 1-366 of the MID1 protein. The $RB_{KAP}$-$CC_{MID}$ is a fusion between the RING finger, B boxes of KAP-1 (residues 22-254) and the coiled-coil region of MID1 (residues 220-366). The $R_{KAP}$-$BCC_{MID}$ is a fusion between the RING finger of KAP-1 (residues 22-138) and the B boxes, and coiled-coil region of MID1 (residues 72-366). The $R_{MID}$-$BCC_{KAP}$ is a fusion between the RING finger of MID1 (residues 1-71) and the B boxes, and coiled-coil region of KAP-1 (residues 139-418). The KAP-RB-[$HSF_{Tri}$], KAP-RB-[$GCN4_{Dim}$], and KAP-RB-[$GCN4_{Tet}$] are heterologous fusions between residues 22-254 of KAP-1 and the trimerization domain of HSF (residues 321-441), the leucine zipper of GCN4 (residues 253-281), and the synthetic tetramerization domain of GCN4 (residues 250-281). A summary of the KRAB binding potential is indicated by plus signs at the right.

KRAB domain binding (Figure 3). Our initial mutational strategy was to target the critical metal chelation residues. For the RING finger and B-boxes, we mutated the potential Zn chelation residues in pairs (or multiples) to lessen the possibility that alternative chelation residues from adjacent

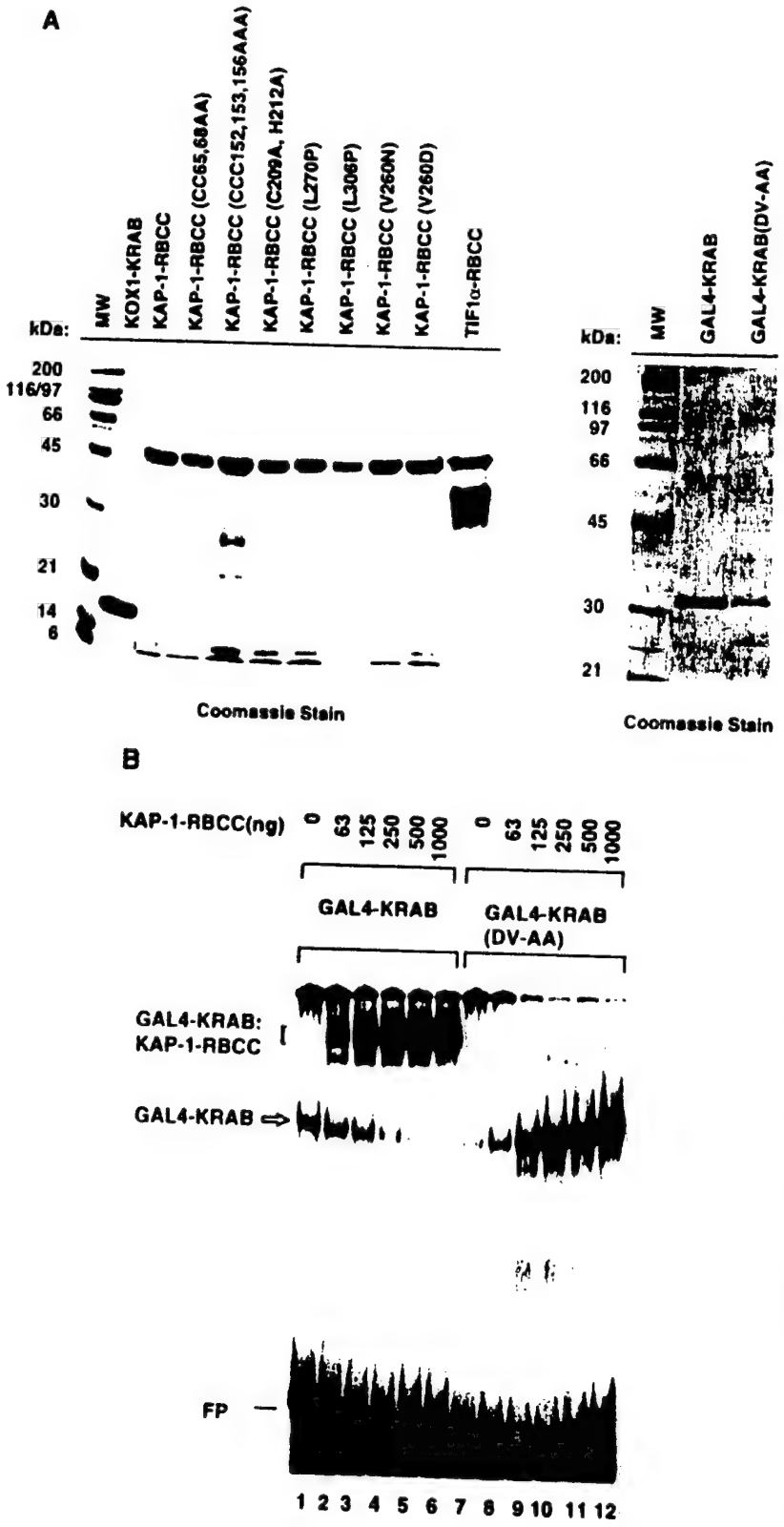



**Figure 2.** Purification, and functional analysis of recombinant proteins expressed in *E. coli* (a) A Coomassie blue-stained SDS/polyacylamide gel of the purified recombinant KOX1-KRAB, KAP-1-RBCC, and GAL4-KRAB fusion proteins. From 3 to 5 μg of the purified KOX1-KRAB, KAP-1-RBCC, and TIF1α-RBCC proteins was loaded per a lane: 1 μg of GAL4-KRAB (wild-type and DV-AA mutant) protein was loaded per a lane. The arrow indicates the recombinant KOX1-KRAB protein. The arrowhead indicates the various KAP-1-RBCC and TIF1α-RBCC proteins. (b) Binding of purified bacterial KAP-1-RBCC to a wild-type KRAB domain but not a mutant form as detected by EMSA. EMSA was performed as described in Materials and Methods. Each reaction contained a constant amount of purified GAL4-KRAB (100 ng), or GAL4-KRAB (DV-AA) (100 ng) (a), a $^{32}$P-labeled DNA probe containing the canonical GAL4 binding site, and increasing amounts of purified bacterial KAP-1-RBCC (a). The DNA:protein complexes were resolved in non-denaturing polyacrylamide gels. The open arrow indicates the DNA:GAL4-KRAB binary complex. The slower migrating DNA:GAL4-KRAB:KAP-1-RBCC complex is represented by the bracket. FP, free probe.

sequence could be recruited to restore metal binding. A double substitution (CC65,68AA) was made at two conserved cysteine residues in the consensus C3HC4 signature (Figure 1(b)) that have been shown to be critical for both coordination of zinc and for stabilization of the hydrophobic core (Saurin *et al.*, 1996). Mutations in the RING finger completely abolished the ability of the KAP-1-RBCC domain to interact with the KRAB domain (Figures 1(d) and 3(a) and (c)).

The second structural element in the RBCC tripartite motif is the B box, which is characterized by a signature spacing of cysteine and histidine residues (Borden, 1998) (Figure 1(b)). Surprisingly, a triple substitution (CCC 152,153,156AAA) in the $NH_2$-terminal B box had no affect on the KAP-1-RBCC binding to the KRAB domain (Figures 1(d) and 3(c)). However, a double substitution (C209A,H212A) in the second B box (Figure 1(b)) strongly reduced the KAP-1-RBCC and KRAB domain interaction (Figures 1(d) and 3(c)).

The coiled-coil domain of KAP-1 is predicted to fold into two potential leucine zipper-like motifs (Wolf *et al.*, 1997) (Figures 1(b) and 3(b)). Each region is likely to be an amphipathic α helix, having a seven-residue repeat $(\mathbf{a}.\mathbf{b}.\mathbf{c}.\mathbf{d}.\mathbf{e}.\mathbf{f}.\mathbf{g})_n$, with hydrophobic residues at positions **a** and **d** and polar residues at the other positions (Figure 3(b)). We altered the leucine zipper motif by substituting a proline residue for a hydrophobic amino acid, which would be predicted to disrupt the α-helix structure. The L270P mutation at position **d** in coil I shows an intermediate effect on the KRAB:KAP-1 interaction (Figures 1(d) and 3(a) and (c)). The L306P mutation at position **d** in coil II was found to completely disrupt the KRAB:KAP-1 interaction (Figures 1(d) and 3(c)). We also tested whether the substitution of a polar amino acid for a hydrophobic amino acid at position **a** would interfere with the KRAB:KAP-1 interaction (Figure 3(b)). However, change of Val260 to either Asn or Asp displayed little effect on the KRAB:KAP-1-RBCC interaction (Figures 1(d) and 3(c)). Together, the above data suggest that an intact RING finger, B boxes, and the coiled-coil domain are each important for the stable binding of the KAP-1-RBCC to a KRAB domain.

Since TIF1α is 40% identical with KAP-1 within the RBCC domain (Figure 1(b)), we investigated whether the RBCC domain of TIF1α could interact with the KRAB domain. We expressed and purified the RBCC domain of TIF1α (residues 1-414) (Figure 2(a)). The RBCC domain of TIF1α displayed no binding to the KRAB domain (Figure 3(c)), an observation that is in contrast with a previous report (Moosmann *et al.*, 1996). It is possible that the KRAB-TIF1α interaction reported by Moosmann was indirect, as it was detected only in yeast two-hybrid assays and was independent of the RING finger and B boxes.

### KAP-1-RBCC exists in an oligomeric state

To investigate the mechanisms behind the KRAB:KAP-1 interaction, we employed biochemical and biophysical analyses of the KRAB domain, the KAP-1-RBCC domain, and the complex. Gel-filtration chromatography was used to estimate the native hydrodynamic size (Stokes radius) of the KAP-1-RBCC protein, which eluted with an apparent molecular mass of 158 kDa (Figure 4(a)-I). This value is roughly three times the molecular mass of the KAP-1-RBCC monomer, suggesting that the KAP-1-RBCC domain is an apparent trimer. This result is consistent with multi-coil scoring analysis (Wolf *et al.*, 1997) for the KAP-1 coiled-coil region.

To determine the oligomeric state of the KAP-1-RBCC in the presence of the native KRAB domain, a KRAB:KAP-1-RBCC complex was pre-formed in solution, and then subjected to gel filtration. This experiment required purified KOX1-KRAB protein from *E. coli*, which initially remained profoundly insoluble. We also attempted to express the KRAB domain from ZNF133, ZNF140, ZNF141, KRK-1, and EEK-1, which also proved insoluble. We finally succeeded in producing soluble KOX1-KRAB protein by employing a protocol that refolds the protein while bound to the Ni-NTA resin (Shi *et al.*, 1997). The KOX-1-KRAB domain was purified to greater than 98% purity and migrated on SDS-PAGE with an apparent molecular mass of 15.5 kDa, which was slightly larger than the predicted molecular mass of 13.9 kDa (Figure 2(a)). Full analysis of this KOX1-KRAB protein will be described elsewhere. However, secondary structure analysis using circular dichroism spectroscopy indicated that the protein was properly folded, and the protein was highly active in binding assays. As shown in Figure 4(a)-III, the KAP-1-RBCC domain in KOX1-KRAB:KAP-1-RBCC complex eluted from the gel-filtration column as a distribution of species, most likely existing in a complex equilibrium, where the majority eluted with an apparent molecular mass consistent with a trimer. Note the slightly larger mass observed for the KAP-1RBCC complexed with the KRAB domain (fraction 11, Figure 4(a)-III) compared to uncomplexed KAP-1-RBCC (fraction 12, Figure 4(a)-I). However, other species exist and these can co-elute with the KRAB domain. The majority of the KOX1-KRAB protein eluted at 670 kDa, suggesting that this protein aggregates when it is not complexed with KAP-1-RBCC (Figure 4(a)-II). To confirm that this complex was stable during chromatography, the peak fraction (11) containing the complex was re-chromatographed under identical conditions. The peak of the KOX1-KRAB:KAP-1-RBCC complex eluted exactly as the previous run, with an apparent molecular mass slightly larger than 158 kDa (Figure 4(a)-IV). We conclude that (1) the KOX1-KRAB:KAP-1-RBCC complex can be readily formed in solution and is stable to gel filtration under these conditions; (2) there is a small but significant increase in the Stokes radius of the complex compared with KAP-1-RBCC alone; and (3) the KAP-1-RBCC domain exists as an oligomer, probably a trimer in the absence or presence of the KRAB domain.

We next used velocity sedimentation in sucrose gradients to estimate the native mass of the KOX1-KRAB:KAP-1-RBCC complex and its components. The KAP-1-RBCC protein alone sedimented with an apparent mass of 66 kDa (Figure 4(b)-I), suggesting either that the oligomeric state of the KAP-1-RBCC in solution was not stable to the conditions of sucrose gradient sedimentation, or that

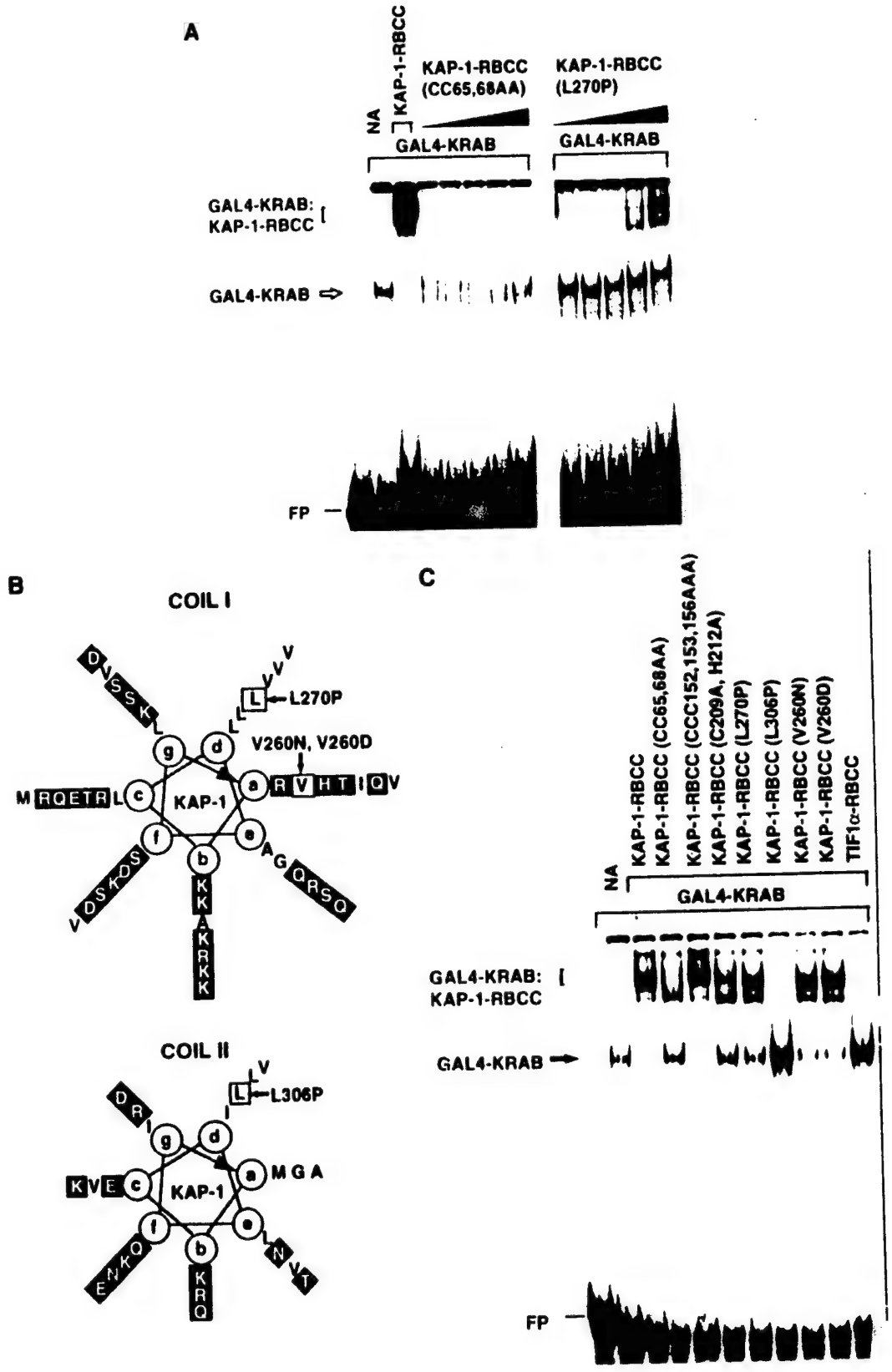



**Figure 3** (*legend opposite*)

the protein was in an extended configuration that results in the relatively large Stokes radius observed by gel filtration. The KOX1-KRAB protein alone sedimented with an apparent masses of 670 and 13.7 kDa (Figure 4(b)-II). However, preformation of a KOX1-KRAB:KAP-1-RBCC complex prior to centrifugation resulted in a dramatic shift in the sedimentation of KAP-1-RBCC domain to an apparent mass of approximately 158 kDa (Figure 4(b)-III). The KOX1-KRAB domain was detected with apparent molecular masses of 670, 158, and 25 kDa, most likely representing aggregated, KAP-1-RBCC-complexed, and monomeric or dimeric uncomplexed KOX1-KRAB protein, respectively. This sedimentation analysis supports the conclusion that the KAP-1-RBCC is oligomeric and that the KRAB domain may stabilize the oligomerization of the KAP-1-RBCC under these sucrose gradient conditions.

We next utilized non-denaturing-PAGE (ND-PAGE) to estimate the native size of KAP-1-RBCC. We calculated the mobilities of the protein standards and of KAP-1-RBCC relative to the tracking dye in a variety of polyacrylamide gel concentrations and used this to generate plots of mobility *versus* polyacrylamide concentration. The slopes of such plots represent the "retardation coefficients" of the proteins (Bryan, 1977). A log-log plot was generated in which the negative retardation coefficients of the standards were plotted against the known native molecular masses of the standards (Figure 4(c)). From this calibration curve, we estimated that the KAP-1-RBCC has a native size of 200 kDa under this condition. These data further substantiate the suggestion that the KAP-1-RBCC protein exists as an oligomer.

To further investigate the oligomeric state of the KAP-1-RBCC domain, sedimentation equilibrium experiments were performed using analytical ultracentrifugation. For each experiment, analyses were performed at 4 °C and 25 °C, and the concentration of protein *versus* radius data was fitted with various models of self-association using non-linear regression (Johnson *et al.*, 1981). The KAP-1-RBCC protein showed non-ideal behavior under these conditions. For two independent experiments, each containing three samples of differing KAP-1-RBCC protein concentration, the data were best described by a model containing monomer, trimer, and hexamer terms (Figure 5). However, when multiple data sets were fitted simultaneously, neither the monomer-trimer or monomer-hexamer association could be described by a single equilibrium constant, indicating that these associations are not reversible on the time-scale of the experiment. The apparent equilibrium constants describing the monomer-trimer and monomer-hexamer association increased with decreasing protein concentration loaded, which suggests that this apparent irreversibility is due to the presence of incompetent trimer and hexamers, i.e. oligomers that do not dissociate. There may also be a reversible association, however, as there was some redistribution of the species (most obviously a reduction in the proportion of hexamer) when the temperature was raised from 4 °C to 25 °C in a single experiment. Unfortunately, we were unable to obtain useful data starting with a pre-formed KOX1-KRAB:KAP-1-RBCC complex, because the protein concentrations of complex that could be obtained were not high enough and a significant portion of the sample formed aggregates under the conditions required for the ultracentrifugation experiments. Nonetheless, these analyses suggest the interpretation that KAP-1-RBCC is predominantly a trimer in solution and a trimer-hexamer equilibrium may play a physiological role.

We next employed chemical cross-linking analyses with ethylene glycol bis (succinimidylsuccinate) (EGS) to determine the oligomeric state of the KAP-1-RBCC proteins (wild-type and mutants) (Figure 6). Treatment of purified KAP-1-RBCC protein with low concentrations of EGS yielded four distinct cross-linked species with apparent molecular masses of 100, 135, 210, and 270-300 kDa, potentially corresponding to dimeric, trimeric, tetrameric, and hexameric forms of the KAP-1-RBCC domain (Figure 6(a)). The monomeric form of the KAP-1-RBCC domain migrated at 49 kDa in SDS-PAGE. Treatment of a pre-formed KOX1-KRAB:-

**Figure 3.** Mutations within the RING finger, B boxes, and the coiled-coil region of the KAP-1-RBCC abolish interaction with the KRAB domain *in vitro*. (a) Binding of the wild-type KAP-1-RBCC or the mutant KAP-1-RBCC (CC65,68AA)/(L270) proteins to the wild-type GAL4-KRAB as detected by EMSA. Wild-type KAP-1-RBCC (1 μg) or increasing amounts of mutant protein (63, 125, 250, 500, 1,000 ng) (Figure 2(a)) were incubated with the wild-type GAL4-KRAB protein. The open arrow indicates DNA:GAL4-KRAB mobility shift. The bracket indicates the DNA:-GAL4-KRAB:KAP-1-RBCC ternary complex. (b) An illustration of the putative heptad repeat for coil I (residues 253-294) and coil II (residues 299-322) of the KAP-1 coiled-coil region. Note the potential amphipathic nature of the putative helices. The heptad positions are labeled by letters **a** through **g** in the helix, where hydrophobic residues at positions **a** and **d** might constitute the cores of homo-trimeric oligomerization domain. The hydrophilic residues at the other positions are indicated with the black background. The two leucine residues (L270 and L306) mutated to proline, and the valine residue (V260) mutated to asparagine or aspartic acid are boxed. (c) The interaction between the KRAB domain and the KAP-1-RBCC domain is dependent upon the integrity of the RBCC domain. The binding of GAL4-KRAB to KAP-1-RBCC or various mutant KAP-1-RBCC proteins as detected by EMSA. A constant amount of GAL4-KRAB (100 ng), 1 μg of each KAP-1-RBCC protein (wild-type or mutant), and 1 μg of TIF1α-RBCC protein was added to each reaction. The arrow indicates the mobility shift for the GAL4-KRAB protein. The bracket indicates the

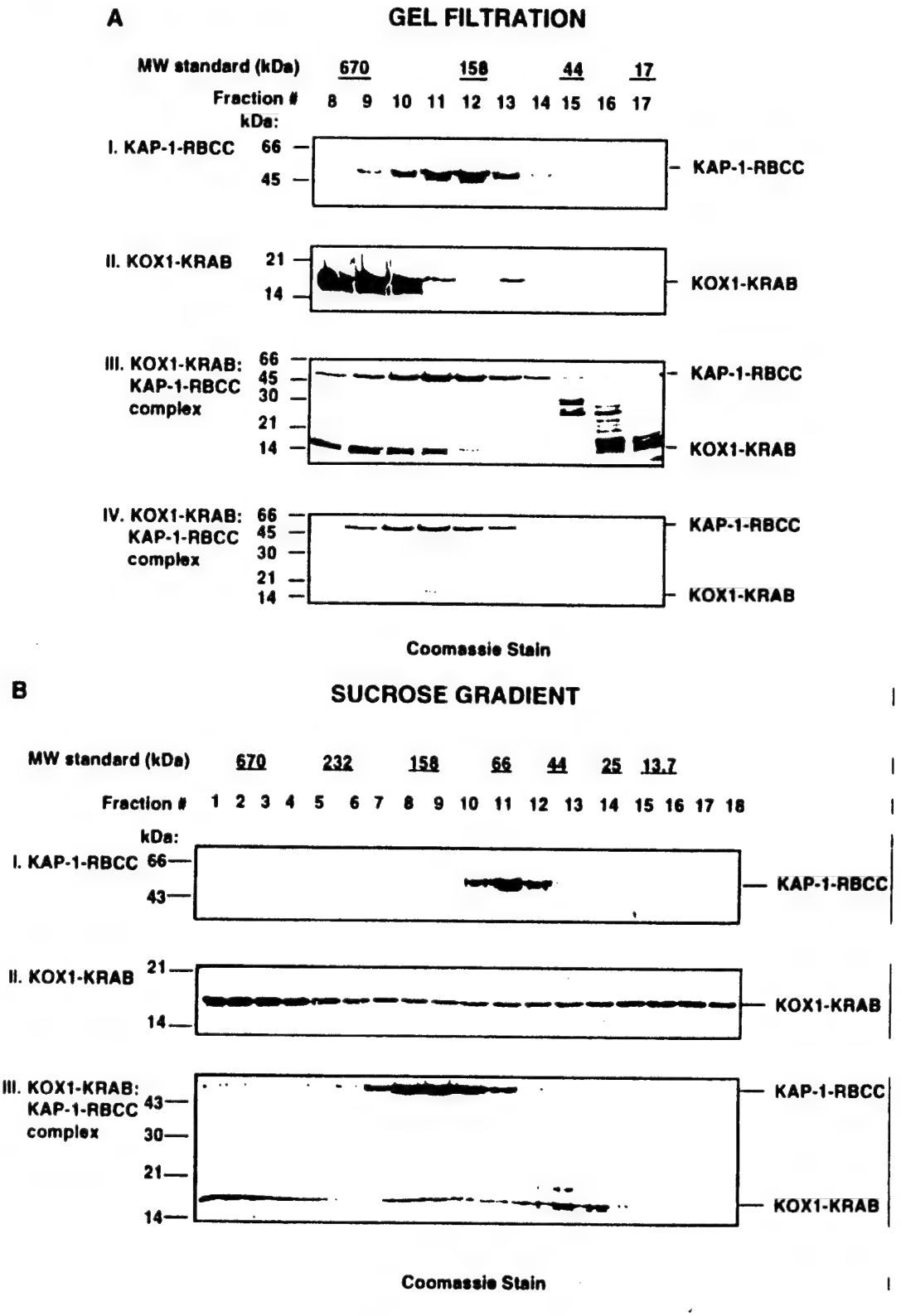



**Figure 4** (*legend opposite*)

KAP-1-RBCC complex with low concentration of EGS yielded the same set of cross-linked species as was observed for the KAP-1-RBCC domain. However, a cross-linked species consistent with the formation of a trimeric KAP-1-RBCC domain appeared as the predominant species even at very

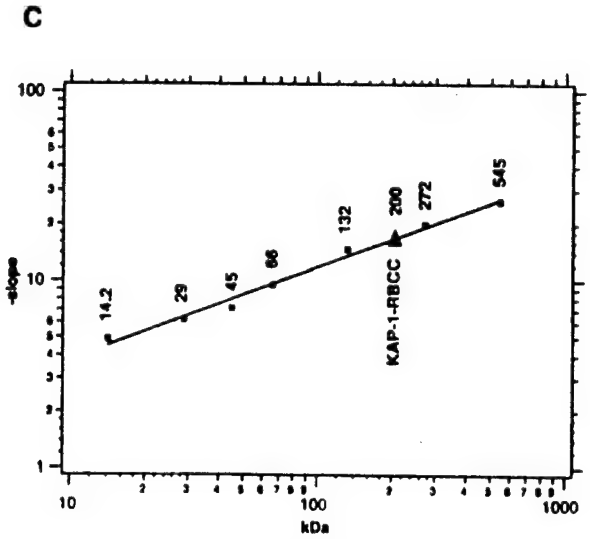



**Figure 4.** Analyses of the KAP-1-RBCC motif, KOX1-KRAB domain, and a KOX1-KRAB:KAP-1-RBCC complex by gel-filtration, sucrose gradient sedimentation, and ND-PAGE. (a) Gel-filtration analysis of KAP-1-RBCC protein, KOX1-KRAB protein, and a pre-formed KOX1-KRAB:KAP-1-RBCC complex. The proteins were resolved on a Superdex 200 column. The fractions were collected and proteins were analyzed by SDS-PAGE. The fraction numbers are labeled at the top of the Figure. The migration of protein standards is indicated above the gels. The protein standards were run in a parallel gel-filtration experiment under identical conditions. (I) Elution profile of native KAP-1-RBCC protein; (II) the elution profile of KOX1-KRAB protein; (III) the elution profile for a pre-formed KOX1-KRAB:KAP-1-RBCC complex; (IV) the profile of the material from fraction 11 (III) re-chromatographed on the column. (b) Sucrose gradient sedimentation analysis of the KAP-1-RBCC domain, KOX1-KRAB domain, and the KOX1-KRAB:KAP-1-RBCC complex. The proteins from sucrose gradient fractions were analyzed by SDS-PAGE. The fraction numbers are indicated at the top of each lane. The peak position of protein standards obtained from gradients run in parallel is indicated above the gels. (c) Determination of the molecular mass of native KAP-1-RBCC by ND-PAGE analysis. The calibration curve prepared from Ferguson plots is shown (see Materials and Methods for details). The known masses of standards and the calculated mass of native KAP-1-RBCC are given (in kDa) above the curve.

low levels of EGS (Figure 6(b)). These data suggest that KRAB domain preferentially binds to a discrete KAP-1-RBCC oligomer and stabilizes this oligomeric form, which is most consistent with a trimer. These cross-linking results are consistent with the sedimentation observed for a pre-formed KOX1-KRAB:KAP-1-RBCC complex in sucrose gradients (Figure 4(b)-III).

Analysis of the mutant KAP-1-RBCC proteins (CC65,68AA) and (L270P), which were severely impaired in KRAB binding (Figures 1(d) and 3(a) and (c)), showed little to no cross-linked species (dimer, trimer, or hexamer) under the same cross-linking conditions (Figure 6(c) and (d)). Furthermore, these mutant proteins migrated with an apparent molecular mass of approximately 66 kDa in sucrose gradients and this migration was not altered by pre-incubation with the KOX1-KRAB protein (data not shown). These data suggest that the formation of the KOX1-KRAB:KAP-1-RBCC complex likely requires the proper oligomeric state of the KAP-1-RBCC, which is apparently a trimer, and that KRAB binding stabilizes this state in solution.

The studies described above were exclusively performed with bacterial-expressed proteins that lack any eukaryotic post-translational modifications and the potential to interact with endogenous cellular partner proteins. To determine if our observations extend to eukaryotic cell-expressed proteins, we overexpressed the KAP-1-RBCC and its mutant derivatives in Sf9 insect cells using a baculovirus expression vector. Full-length wild-type KAP-1, KAP-1-RBCC and mutants thereof were highly expressed as soluble proteins in Sf9 cells. The bv.KAP-1-RBCC was purified under native conditions and migrated as a single polypeptide band consistent with its predicted monomeric molecular mass of 47.7 kDa (data not shown). This behavior suggests that bv.KAP-1-RBCC is not significantly subjected to extensive post-translational modification, as detected by SDS-PAGE when expressed in eukaryotic cells. The bv.KAP-1-RBCC protein behaved identically with the *E. coli*-expressed protein in all of the biochemical binding assays and was equally active in KRAB domain interaction (data not shown).

To estimate the native mass and oligomeric states of KAP-1 proteins (wild-type and mutants) expressed in the baculovirus expression system, lysates from infected cells that contain the expressed proteins were subjected to gel-filtration chromatography and detected in a Western blot assay (Figure 7(a)). The full-length KAP-1 protein eluted with an apparent molecular mass of 500 to 670 kDa (Figure 7(a)-I). This elution profile was

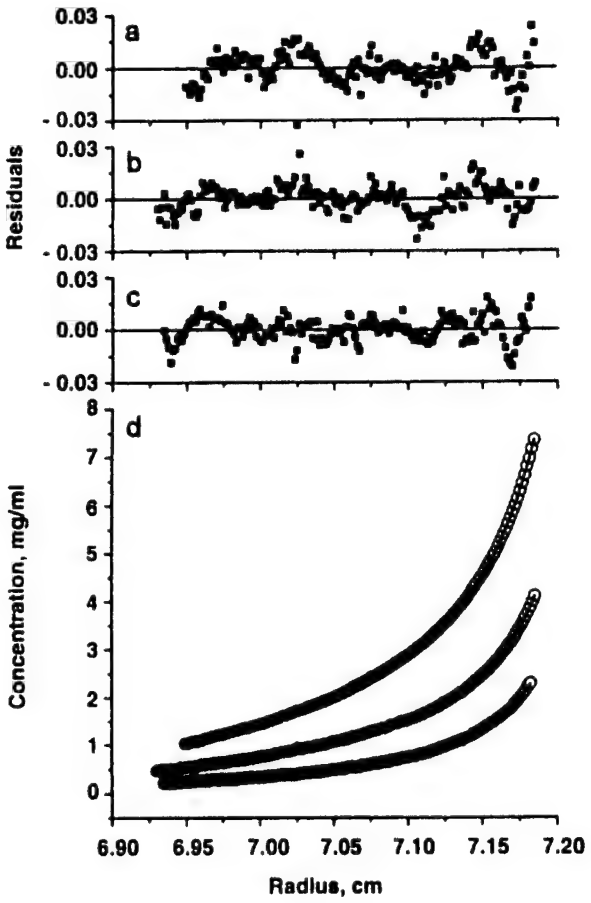



**Figure 5.** Representation of analytical ultracentrifugation analysis of KAP-1-RBCC. Sedimentation equilibrium analysis of KAP-1-RBCC was performed with three different loading concentrations in separate cells at 13,000 rpm and 4°C. The circles in the main panel (d) show the concentration *versus* radius data for the three cells at equilibrium. The three data sets were fitted separately with a model describing monomers, trimers, and hexamers, using the non-linear regression program NONLIN (Johnson *et al.*, 1981). The calculated fit is represented by the continuous lines. (a)-(c) show the residuals for the data points of the fitted curves at the three protein concentrations displayed from highest concentration (top) to lowest concentration.

identical with both bacterial-expressed full-length KAP-1 and endogeneous KAP-1 derived from mammalian cell nuclear extracts (unpublished results). Since the apparent monomeric molecular mass measured by SDS-PAGE of full-length KAP-1 is 97 kDa, this result suggests that native full-length KAP-1 was an oligomer, possibly as large as a hexamer. The bv.KAP-1-RBCC protein co-eluted with the *E. coli* KAP-1-RBCC protein at approximately 158 kDa (Figure 7(a)-II), suggesting that *E. coli* and bv.-expressed KAP-1-RBCC share a common oligomeric state. Mutations in the RING finger (CC65,68AA) and the coiled-coil region (L270P) of the KAP-1-RBCC appeared to alter oligomerization properties, as these mutant proteins eluted with an apparent molecular mass of 100 kDa (Figure 7(a)-III,IV), and also failed to bind GAL4-KRAB in the EMSA assay (Figure 8(a), lanes 3 and 4). These data are consistent with those observed for proteins produced in *E. coli*, and strengthen the argument that KAP-1-RBCC must properly oligomerize for KRAB binding, which is independent of post-translational modifications.

### Formation and detection of KAP-1:KAP-1-RBCC hetero-oligomers *in vivo*

To reconstitute hetero-oligomers of the KAP-1-RBCC domain with full-length KAP-1 *in vivo*, we co-expressed full-length KAP-1 with either wild-type or mutant KAP-1-RBCC proteins *via* co-infection of Sf9 cell cultures (Figure 7(b)). Because the KAP-1-RBCC encodes an $NH_2$-terminal 6His tag, its association with wild-type full-length KAP-1 could be detected by Ni-NTA chromatography, followed by Western blot analysis using an anti-KAP-1 serum that detects amino acid residues 428-589, a region that is not present in the KAP-1-RBCC protein (Friedman *et al.*, 1996) (Figure 7(b)). The KAP-1-RBCC bound to the Ni-NTA resin was detected with an anti-His monoclonal antibody. The full-length bv.KAP-1 (which has no 6His tag) did not bind to Ni-NTA resin (Figure 7(b), lane 1). When the Ni-NTA resin was incubated with Sf9 cell lysates, in which full-length KAP-1 protein was co-expressed with the KAP-1-RBCC, we observed retention of the full-length KAP-1 protein (Figure 7(b), lane 2). These results suggest that the KAP-1-RBCC domain can hetero-oligomerize with the full-length KAP-1 protein. Interestingly, simple mixing of cell lysates that contained independently expressed proteins failed to recapitulate this interaction (data not shown). Thus, the formation of hetero-oligomers between full-length KAP-1 and the KAP-1-RBCC domain may be coupled to translation or other post-translation processes, an observation that has been made for other oligomeric coiled-coil proteins. The KAP-1-RBCC mutants CC65,68AA and L270P also efficiently associated with full-length KAP-1 *in vivo* using the co-infection assay (Figure 7(b), lanes 3 and 4). This result suggests that these mutations inhibit the homo-oligomerization of the RBCC domain but that these RBCC mutants can still hetero-oligomerize with a wild-type RBCC domain *in vivo*.

Since the CC65,68AA and L270P mutant RBCC proteins could apparently hetero-oligomerize with full-length KAP-1, based upon the data above, we tested the ability of these hetero-oligomers to interact with the KRAB domain in the EMSA assay. The bv.KAP-1-RBCC formed the characteristic complex with the DNA-bound GAL4-KRAB protein (Figure 8(a), lane 2). The mutant proteins CC65,68AA and L270P each failed to form this ternary complex (Figure 8(a), lanes 3 and 4), a result that is consistent with our previous findings, in which *E. coli*-derived mutant RBCC proteins

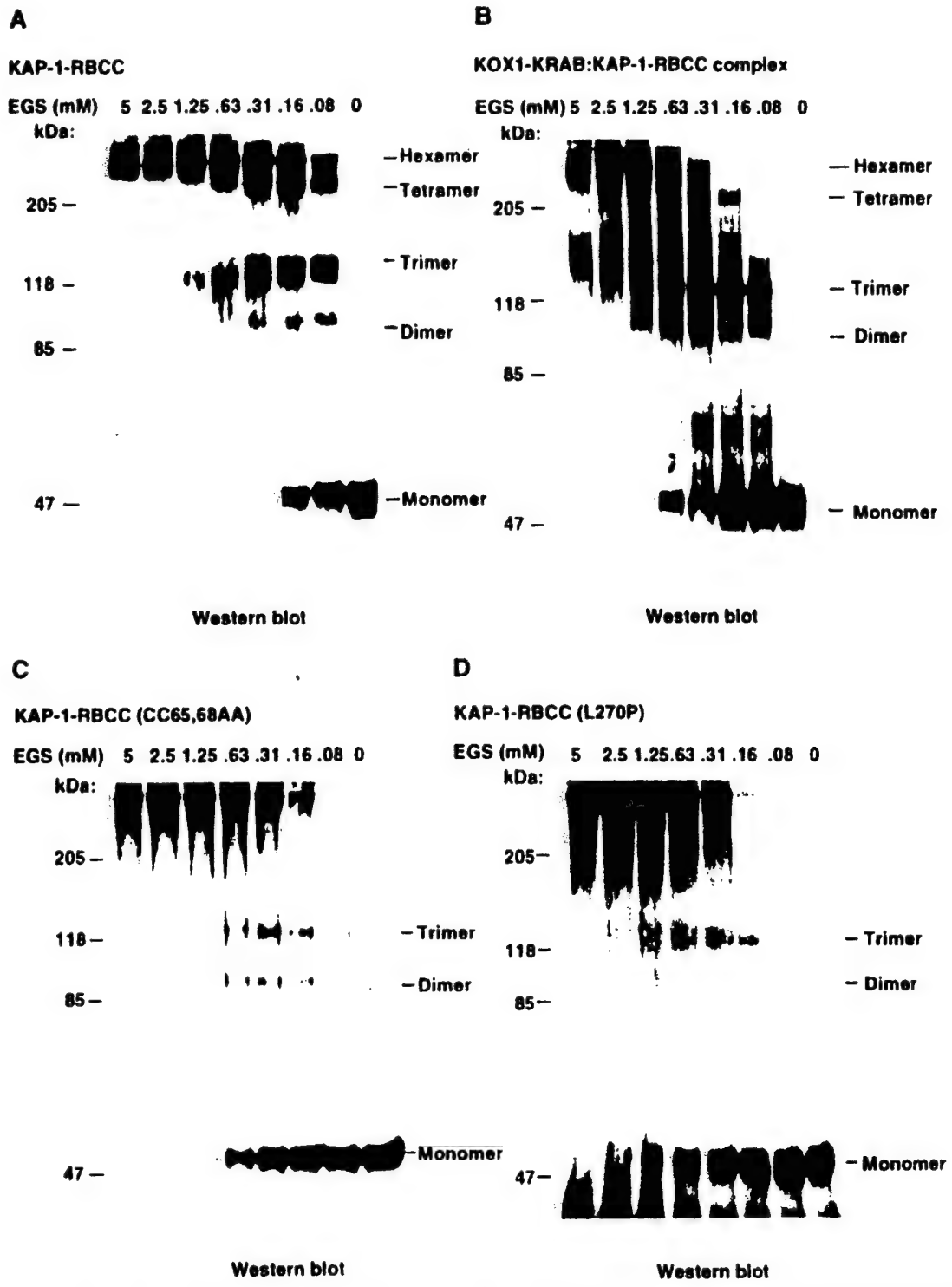



**Figure 6.** Chemical cross-linking analysis of the KAP-1-RBCC domain. Equal amounts (5 μM) of (a) wild-type KAP-1-RBCC, (b) a KOX1-KRAB:KAP-1-RBCC complex, (c) the (CC65,68AA) mutant, and (d) the (L270P) mutant were incubated with various concentrations of EGS. The cross-linked KAP-1-RBCC products were resolved by SDS-PAGE and visualized by Western blot analysis. An anti-His monoclonal antibody (Qiagen) was used to detect the KAP-1-RBCC proteins in panels (a), (c) and (d). An anti-KAP-1-RBCC rabbit IgG was used to detect the KAP-1-RBCC protein in (b). The position of molecular markers is indicated on the left of each Western blot. The estimated multimeric forms of the proteins are indicated on the right of each Western blot.

failed to bind the KRAB domain (Figures 1(d) and 3(a) and (c)). The full-length bv.KAP-1 protein yielded a supershift that was easily distinguishable from the supershift produced by bv.KAP-1-RBCC (Figure 8(a), lane 5). It is well established that this difference can be exploited as an assay for hetero-

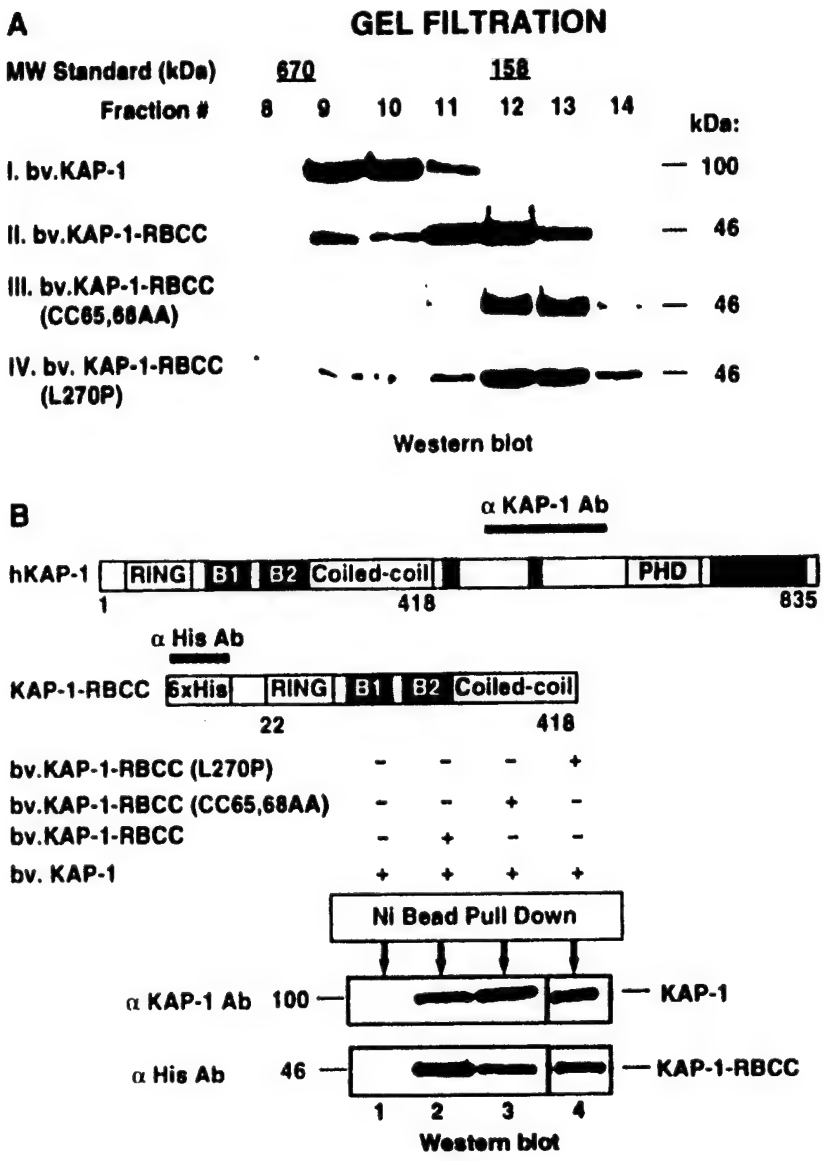



**Figure 7.** Expression and functional analysis of KAP-1 proteins expressed from baculovirus vectors. (a) Gel-filtration analysis of full-length KAP-1 (I), wild-type KAP-1-RBCC (II), and mutant KAP-1-RBCC (III, IV). The elution of each protein from a Superdex 200 column was detected by Western blot analysis. The fraction numbers are labeled at the top of each lane, and the protein standards are indicated above the blots. An affinity-purified rabbit polyclonal anti-KAP-1 serum was used to detect full-length KAP-1. An anti-His monoclonal antibody was used to detect the KAP-1-RBCC proteins. (b) Formation and detection of KAP-1:KAP-1-RBCC hetero-oligomers *in vivo*. Diagrams of full-length KAP-1 and 6His KAP-1-RBCC proteins expressed in baculovirus are presented. The proteins were either expressed individually, or co-expressed in Sf9 cells. Hetero-oligomerization of the KAP-1-RBCC domain with full-length KAP-1 was detected by binding the complex to Ni-NTA resin. The bound proteins were detected by Western blot analysis. The bound full-length KAP-1 protein was detected using a rabbit polyclonal anti-KAP-1 serum, and the KAP-1-RBCC proteins were detected with the anti-His monoclonal antibody. The + represents the viruses used to infect Sf9 cells.

oligomerization of proteins by searching for intermediate mobility supershifts (Hope & Struhl, 1987). The bv.KAP-1 and bv. KAP-1-RBCC co-infected extracts produced a spectrum of five intermediate supershifts (Figure 8(a), lane 6), including the bands that co-migrate with the authentic bv.KAP-1 and bv.KAP-1-RBCC supershifts. This suggests that discrete hetero-oligomers composed of various combinations of full-length KAP-1 and wild-type RBCC domain can bind to the KRAB domain. Each of these supershifts contains either KAP-1, KAP-1-RBCC, or a combination thereof as determined by antibody supershift experiments (data not shown). If the complex between bv.KAP-1 and bv.KAP-1-RBCC was a simple hetero-dimer, a single intermediate species would be predicted. The presence of multiple discrete species strongly suggests that hetero-oligomers are formed, most likely hetero-trimers. Co-expression of KAP-1 with either the mutant RBCC protein yielded mobility shifts that were consistent only with the supershift observed for full-length KAP-1 (Figure 8(a), lanes 7 and 8). A plausible interpretation of these results is illustrated by Figure 8(b). Bands A and B represent hexamer and trimer forms of full-length KAP-1, respectively, which are in equilibrium; the hexamer is formed most efficiently by the full-length KAP-1 protein and binds with highest affinity to the KRAB domain. Bands C and D are hetero-trimers that contain one and two molecules of KAP-1-RBCC per complex, respectively. These complexes can bind the KRAB domain but may lack the ability to form a hexamer under these conditions. Band E is the homo-trimer of wild-type KAP-1-RBCC. The mutant RBCC proteins retain the ability to hetero-trimerize, but any complex containing a single mutant molecule would be inactive for KRAB domain binding, as evidenced by the loss of complexes C, D, and E in Figure 8(a), lanes 7 and 8. These results suggest that oligomerization alone

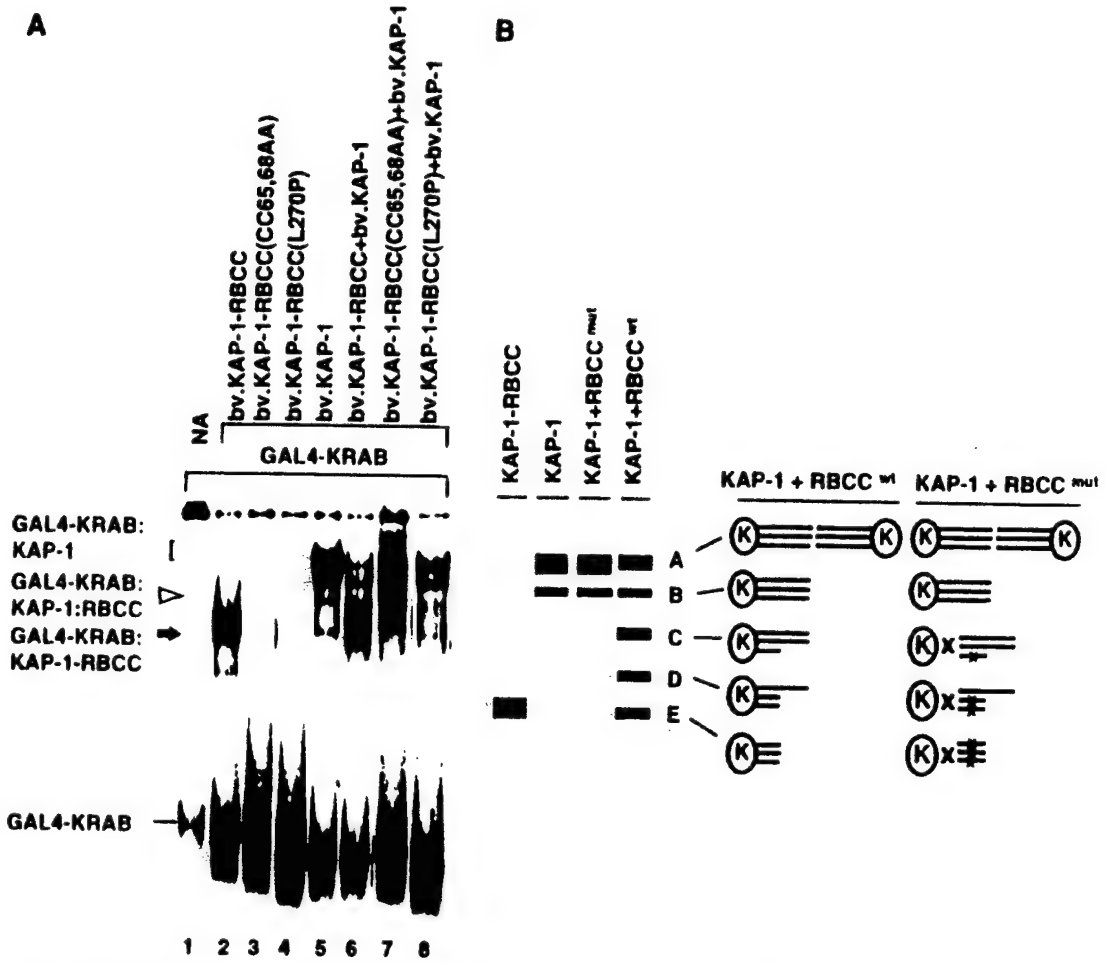



**Figure 8.** Binding of KAP-1-RBCC:KAP-1 hetero-oligomers to a DNA-bound KRAB domain *in vitro*. (a) The binding of KAP-1-RBCC:KAP-1 hetero-oligomers to GAL4-KRAB as detected by EMSA. Each reaction contained a constant amount of purified GAL4-KRAB (100 ng), 5 μg of baculovirus whole-cell lysates, in which full-length KAP-1 protein was co-expressed with the wild-type or mutant KAP-1-RBCC, and 1 μl of $^{32}$P-labeled Gal4 probe. The symbols on the left side of the panel defines the following complexes: dash, DNA:GAL4-KRAB alone; filled arrow, DNA:GAL4-KRAB:KAP-1-RBCC complex; open arrowhead, DNA:GAL4-KRAB:KAP-1:RBCC complex; bracket, DNA:GAL4-KRAB:KAP-1 complex. The free probe was allowed to run off the bottom of the gel for better resolution of the mobility shifts at the top of the gel. Note that no binding of the (CC65,68AA), or the (L270P) mutant KAP-1-RBCC proteins to GAL4-KRAB is detected (lanes 3 and 4). Also, hetero-oligomers of full-length KAP-1 with either of the mutant KAP-1-RBCC proteins failed to yield any intermediate supershifts (lanes 7 and 8) of the GAL4-KRAB complex, which were observed for hetero-oligomers of full-length with the wild-type KAP-1-RBCC domain (lane 6). (b) An interpretation for the formation of GAL4-KRAB:KAP-1:KAP-1-RBCC complexes in the EMSA assay in (a).

is not sufficient for interaction of KAP-1 with the KRAB domain. Moreover, individual components within the RING finger, B boxes, and coiled-coil region independently assist the proper oligomerization of KAP-1 and its interaction with the KRAB domain.

### Heterologous oligomerization domains fail to reconstitute the KOX1-KRAB:KAP-1 interaction

The data presented above suggest a simple model for the KAP-1-RBCC domain: that the coiled-coil region is a parallel trimeric association that serves to properly orient the RING and B boxes for binding to the KRAB domain. If this hypothesis is correct, we may be able to substitute a similarly structured coiled-coil domain to reconstitute KRAB binding. We thus substituted the coiled-coil region of KAP-1 with well-characterized heterologous oligomerization domains for trimerization (HSF) (Drees *et al.*, 1997), dimerization (GCN4-LZ), and tetramerization (GCN4-TZ) (Waterman *et al.*, 1996) (Figure 1(d)). These heterologous KAP-1-RBCC fusion proteins were expressed as soluble proteins and were properly oligomerized, as determined by chemical cross-linking and gel filtration (data not shown). As illustrated in Figures 1(d) and 9, each of these proteins, KAP-RB-[$HSF_{Tri}$], KAP-RB-[$GCN4_{Dim}$], and KAP-RB-[$GCN4_{Tet}$], failed to interact with the DNA-bound GAL4-KRAB protein. Thus, the KAP-1 coiled-coil region may provide a unique oligo-

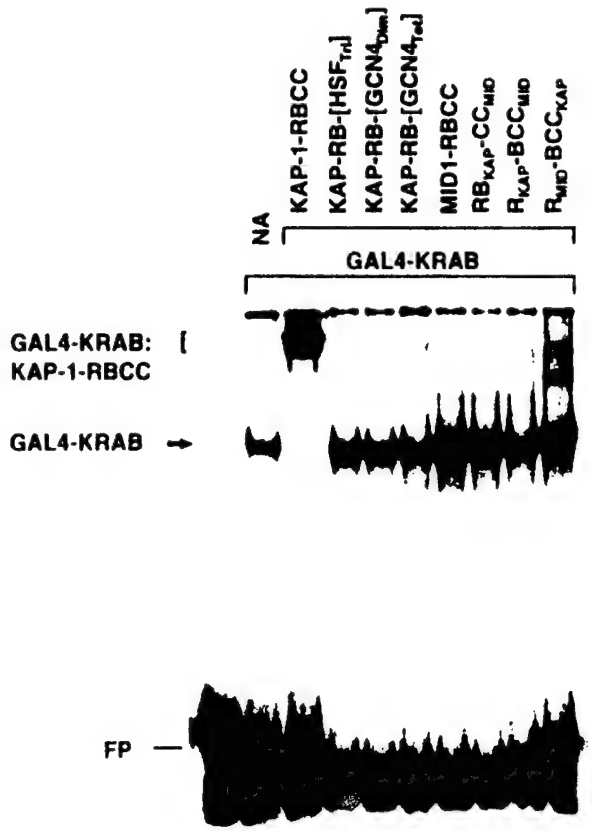



**Figure 9.** Substitution of the RING finger, B-boxes, and the coiled-coil sub-domains of KAP-1 disrupts the KOX1-KRAB:KAP-1 interaction: 1 μg of each wild-type KAP-1-RBCC or heterologous KAP-1-RBCC fusion protein was tested for its ability to interact with the KRAB domain *in vitro*, *via* the EMSA. The wild-type KAP-1-RBCC protein efficiently supershifted the DNA-bound GAL4-KRAB protein to form a complex indicated by the bracket. None of the heterologous RBCC fusion proteins demonstrated this activity, although the $R_{MID}BCC_{KAP}$ fusion protein showed a little of this activity. The filled arrow indicates the DNA:GAL4-KRAB mobility shift. FP, free probe.

merization surface required for KRAB binding or contribute directly to KOX1-KRAB domain recognition *via* specific contacts.

The studies described here suggest that the RING finger, B boxes, and coiled-coil region collectively appear to constitute a tripartite interface for protein-protein interactions that is conserved in a number of cellular proteins including TIF1α, TIF1γ, and MID1 (Figure 1). To more precisely address the specificity of each motif in this apparent protein-protein interaction domain of KAP-1, we have substituted the RING finger, B boxes, and coiled-coil sub-domains of KAP-1 with the corresponding motif from MID1 (Quaderi *et al.*, 1997) (Figure 1(d)). We expressed and purified the MID1-RBCC domain and heterologous KAP-1/MID1 RBCC fusion proteins from bacteria (data not shown), and tested their ability to interact with the KRAB domain. In general, the MID1-RBCC domain and each of the KAP-1/MID RBCC fusion proteins failed to bind to the DNA-bound GAL4-KRAB protein. However, $R_{MID}BCC_{KAP}$ fusion protein showed weak potential to shift the DNA:-GAL4-KRAB binary complex (Figures 1(d) and 9).

Together, these data suggest that the RBCC domain of KAP-1 functions as an integrated structural unit to independently facilitate the oligomerization of KAP-1 and its highly specific interaction with the KRAB domain.

## Discussion

The RBCC tripartite domain has emerged as a highly conserved module in proteins that play causal roles in cellular differentiation and development, and whose alteration can result in human disease (El-Husseini & Vincent, 1999; Ogawa *et al.*, 1998; Perrin & Lacroix, 1998; Tatematsu *et al.*, 1998). The spatial conservation of sub-domains in the RBCC motif and their presence in large cellular complexes has led to the suggestion that the RBCC domain mediates specific protein-protein interactions. However, these targets, even for the most well-characterized RBCC proteins have eluded investigators.

Here, we present the first demonstration that a highly specific, physiologically relevant protein-protein interaction can be reconstituted with a purified RBCC domain and its corresponding binding partner, thus providing proof that the RBCC domain is a protein-protein binding surface. We can draw the following conclusions from our data: (1) the interaction is direct; it does not apparently require post-translational modification and is stable to *in vitro* biochemical manipulation. (2) A stable ternary complex can be formed between DNA, a DNA-bound KRAB domain and the KAP-1-RBCC protein; thus, the KRAB-KAP-1 interaction does not inhibit DNA-binding in the system we have used. (3) Recognition of the KRAB domain by the KAP-1-RBCC is highly specific, as the related RBCC proteins MID1 and TIF1α do not bind under these experimental conditions. (4) Binding requires structural elements of all the sub-domains of the KAP-1-RBCC, as mutation of each individual sub-domain has the potential to abolish KRAB domain binding. (5) Each RBCC sub-domain may co-operate in providing specificity for interaction, as individual sub-domain swaps with related RBCC proteins fail to recapitulate binding. (6) The KAP-1-RBCC domain must properly homo-oligomerize in order to bind the KRAB domain: this oligomerization state is probably a trimer and further association to a hexamer may occur. (7) Oligomerization is necessary but not sufficient for KRAB domain binding, implying that separate functional domains for oligomerization and KRAB recognition exist. (8) The coiled-coil sub-domain itself provides critical binding and/or determines orientational specificity for KRAB domain binding, as

heterologous oligomerization domains fail to recapitulate binding. The picture that emerges is that the KAP-1-RBCC domain forms an integrated unit composed of essential sub-domains that probably function cooperatively to recognize a protein ligand with high specificity.

Although it has been demonstrated that many RBCC proteins oligomerize, its role in ligand binding has not been determined. Our data with the KRAB-KAP-1 system firmly establish that oligomerization is required for binding. A major issue is the oligomerization of KAP-1, before and after KRAB binding and whether binding stabilizes the oligomeric species. We favor the interpretation that the KRAB-bound KAP-1-RBCC domain is a trimer or hexamer based upon a number of observations. Analysis of the predicted coiled-coil region using the multi-coils program predicts the potential for two independent coiled regions with the highest probability that each forms a trimer. Gel filtration of soluble *E. coli* KAP-1-RBCC in the presence or absence of the KRAB domain shows a complex of apparent molecular mass three times greater than the monomeric molecular mass. Chemical cross-linking in solution shows apparent trimeric and hexameric forms of the KAP-1-RBCC as the predominant species, and the production of the trimeric cross-linkable form is greatly enhanced by pre-incubation with recombinant KRAB domain. Analytical ultracentrifugation experiments show mixtures of monomers, trimers, and hexamers. Together, these observations are consistent with the potential for a monomer-trimer-hexamer equilibrium. The KAP-1-RBCC protein also sedimented as an apparent monomer in standard velocity sedimentation in sucrose gradients, suggesting that oligomers were not stable under these conditions. However, pre-incubation of the KAP-1-RBCC with the KRAB domain resulted in a sucrose gradient sedimentation profile of the complex consistent with a trimer. The shift from apparent monomer to apparent trimer in sedimentation is unlikely to be accounted for by simple incorporation of multiple molecules of KRAB domain into the complex, since preliminary stoichiometry estimates suggest that the KAP-1-RBCC:KRAB complex is approximately 3:1 (unpublished results). In summary, the behavior of the complex in these assays most likely indicates KRAB-induced trimer formation or a conformational change detected by sedimentation.

We confirmed these studies using KAP-1-RBCC and full-length KAP-1 protein expressed in eukaryotic cells. The co-infection experiments clearly show that hetero-oligomers can be efficiently formed between full-length KAP-1 and KAP-1-RBCC, and that the hetero-oligomers are functional for KRAB binding. Most importantly, three intermediate supershifts were observed. This strongly suggests that the active binding species is oligomeric, most likely a trimer or hexamer.

The mutations in either the RING finger (CC65,68AA) or the coiled-coil region (L270P) when expressed individually in Sf9 cells were inactive in KRAB domain binding. Thus, either they are defective in homo-oligomerization or form an oligomer unproductive for KRAB binding. Both of these mutant proteins show slightly smaller apparent molecular masses in gel filtration, suggesting an unfolded state. This apparent monomeric species cannot be converted to the apparent oligomeric species (as detected by sucrose gradient sedimentation or cross-linking) *via* incubation with the KRAB domain. Thus, the apparent KRAB-induced oligomerization is required for binding and both the RING finger and the first coiled-coil are important for this oligomerization.

However, it is puzzling that these mutants, when co-expressed with full-length KAP-1 in baculovirus-infected cells efficiently form a hetero-oligomeric complex. Clearly, these hetero-oligomeric complexes are not competent for KRAB domain binding, since there were no intermediate supershift bands observed from a mutant/wild-type co-infected extract. Thus, either the binding is non-specific, or the full-length KAP-1 protein compensates for an oligomerization defect in the mutant RBCC protein.

Hetero-oligomers that contain one mutant subunit are apparently defective in KRAB binding and thus these hetero-oligomers appear to act as true dominant-negatives. This behavior has been observed in many other oligomeric proteins, where functionally-separate oligomerization and ligand binding domains are present (Hurst, 1995; Baxevanis & Vinson, 1993). An important point is that the RING finger mutant (CC65,68AA) fails to form a productive homo-oligomer or hetero-oligomer with wild-type KAP-1, even though it contains an intact coiled-coil region. This observation clearly demonstrates that the RING finger participates both in oligomerization and ligand recognition.

Our mutational analysis of the individual RBCC sub-domains clearly shows that the core structural integrity of the RING finger appears to be required for KRAB domain binding. Many engineered and naturally occurring RING finger mutations have been studied in other proteins but none has yet been correlated with alteration of a specific RING-protein interaction in a purified system. Homo or hetero-oligomerization *via* RING-RING interaction has been observed in PML, BRCA1, Bmi-1, Mph2 (Borden *et al.*, 1995a; Hemenway *et al.*, 1998; Meza *et al.*, 1999). The RING finger mutant BRCA1 proteins show altered oligomerization, subcellular localization, interaction with protein partners, and ability to suppress cell growth (reviewed by Irminger-Finger *et al.*, 1999). RING finger mutations in the proto-oncogene PML have profound effects on the localization of the protein to the ND10 nuclear bodies (Borden *et al.*, 1995a), which have been correlated with effects on both cell survival mediated by wild-type PML (Borden *et al.*, 1997) and oncogenesis as mediated by the PML-RARα fusion (Fagioli *et al.*, 1998). Finally, the RING finger of RFP is critical to its cell transform-

ation function as the rfp-ret fusion oncogene (Cao *et al.*, 1997, 1998; Hasegawa *et al.*, 1996).

The mutational analysis of the B1 and B2 boxes of KAP-1 has also been informative. Interestingly, the triple-cysteine substitution in the $NH_2$-terminal B1 box had no effect on KRAB domain binding, whereas a similar substitution in the analogous cysteine residues in the B2 box completely abolished KRAB domain binding. If these substitutions abolished the structural integrity of both B boxes, it is reasonable to conclude that there is an unequal contribution of each B-box to KRAB domain binding and/or KAP-1 oligomerization. The evidence that the B-box is required for the proper orientation of the coiled-coil region (Cao *et al.*, 1997) could explain why mutation of the B2 box adjacent to the coiled-coil in KAP-1 abolished KRAB binding, whereas mutation of the distal B1 box had no effect.

Oligomerization mediated by the coiled-coil region is one of the most important determinants of the biological properties of RBCC proteins, and is likely responsible for their legendary potential for aggregation and multimerization *in vitro*. The L270P substitution in the KAP-1-RBCC expressed in *E. coli* showed aberrant sedimentation and failure to oligomerize in the chemical cross-linking assay (in the absence of the KRAB domain). These properties correlated with an impaired ability to bind GAL4-KRAB, although significant binding was detected at very high levels of input KAP-1-RBCC (L270P). However, baculovirus-expressed KAP-1-RBCC (L270P) was completely inactive for GAL4-KRAB binding, even at very high concentrations. Surprisingly, the bv.KAP-1-RBCC (L270P) still efficiently bound to wild-type full-length bv.KAP-1 in co-infected cells. However, these hetero-oligomeric complexes were not able to bind the KRAB domain, thus exhibiting an apparent dominant-negative activity essentially identical with the KAP-1-RBCC, RING finger mutant (CC65,68AA) discussed above. Like the RING finger mutant, we see a difference in homo-oligomeric *versus* hetero-oligomeric potential with wild-type KAP-1 in the KAP-1-RBCC (L270P) mutant. The L306P substitution in coil II is clearly different. The *E. coli*-produced L306P protein is completely blocked in its ability to bind the KRAB domain, even at high concentration. Thus, either the L270 and L306 leucine residues that we targeted are not equivalent in terms of their contribution to each coiled-coil structure, or coil I and coil II are not equivalent in their contribution to either oligomerization or KRAB domain recognition. It is a distinct possibility that one (or both) coils participate directly in KRAB domain recognition. In summary, our mutational analysis for the KAP-1-RBCC subdomains strongly suggests that each sub-domain contributes to KRAB domain binding, either as a direct recognition element or indirectly by mediating appropriate homo-oligomerization.

To begin distinguishing these possibilities, we constructed a number of fusion proteins containing heterologous oligomerization domains and "swapped" RBCC sub-domains from KAP-1 orthologues. Our strategy of replacing the coiled-coil region of KAP-1 with heterologous oligomerization domains was based on the following plausible model for the KAP-1-RBCC oligomer: that the coiled-coil region forms a parallel trimeric association of helices that orient the the RING and B-boxes for KRAB binding. If the coiled-coil serves only as an oligomerization motif, it would be predicted that it could be substituted for by a similar motif, as has been shown in other oligomeric proteins. The best evidence that supports such a potential model comes from recent structural analyses of the TRAF proteins (McWhirter *et al.*, 1999; Park *et al.*, 1999), which encode RING finger and coiled-coil regions and are a subfamily of the RBCC proteins. The TRAFs homo and hetero-trimerize *via* a parallel association of the coiled-coil region. Trimerization properly orients the TRAF domain for binding to the cytoplasmic receptor, and this trimerization function can be substituted by heterologous trimerization domains (McWhirter *et al.*, 1999; Park *et al.*, 1999). However, KAP-1-RB proteins fused to heterologous dimerization, trimerization, or tetramerization domains completely failed to bind the KRAB domain. Our alternative approach, using domain swaps from the MID1-RBCC protein yielded similar, uniformly negative results. That is, substitution of the RING, B-box-coiled-coil, or coiled-coil region alone from MID1 for the corresponding region in KAP-1 failed to recapitulate KRAB domain binding. Together, these results with chimeric RBCC proteins strongly suggest that the RBCC domain of KAP-1 binds as an integrated functional unit to its ligand, and that each sub-domain contributes to protein recognition and/or oligomerization. It is noteworthy that the intact RBCC domains from three independent RBCC proteins, TIF1α, MID1 and TIF1γ (Venturini *et al.*, 1999) all fail to bind the KRAB domain. These observations underscore the apparent high degree of specificity for ligand recognition that is built into the RBCC domain of KAP-1 and probably other RBCC proteins.

An important point is whether endogenous RBCC proteins form hetero-oligomers *via* RBCC-RBCC interactions. These interactions would serve to increase the combinatorial complexity for target recognition by hetero-oligomeric RBCC complexes. It is interesting to note that apparent hetero-trimerization has been observed among certain members of the TRAF subfamily of RBCC proteins, although it is unclear whether this alters the biological signaling properties/specificities of these molecules (Pullen *et al.*, 1999). Even among the clearly paralogous TIF1 family of RBCC proteins, only indirect weak binding between family members has been detected (Venturini *et al.*, 1999). Rather remarkably, the MID1 and MID2 RBCC proteins, which show overall 77% amino acid sequence identity fail to

hetero-dimerize *in vivo* (Buchner *et al.*, 1999; Cainarca *et al.*, 1999). However, there is indirect evidence that RFP can hetero-oligomerize and co-localize with PML, and that this activity requires the B box and coiled-coil regions of RFP (Cao *et al.*, 1998). Clearly, studies with purified components as demonstrated here, are required to resolve this important issue.

Our primary goal in this study was to begin to establish a set of principles for oligomerization and ligand recognition for the RBCC family of proteins using the KRAB-KAP-1 system as a model. However, equally important are the implications that our findings have for defining the biological functions of KAP-1 as a corepressor *in vivo*. The KRAB domain is currently represented in more than 100 proteins, most of which are $C_2H_2$ zinc-finger proteins and are likely to function as gene-specific repressors. Although we have exclusively utilized the KRAB domain of KOX1, more than ten KRAB domain proteins have been shown to bind KAP-1 and thus it is a reasonable assumption that the vast majority of this family will utilize KAP-1 as their corepressor. Our current working model for KRAB-ZFP-KAP-1-mediated repression is illustrated in Figure 10. A KRAB-ZFP binds DNA sequence-specifically through its array of $C_2H_2$ zinc fingers. The DNA-bound KRAB domain recruits the KAP-1 corepressor *via* direct interaction with the RBCC domain. Binding may induce oligomerization of KAP-1, which is required or KAP-1 may be pre-existing in an oligomeric state. Binding of KAP-1 by the KRAB-ZFP in proximity to the transcription start site is necessary but not sufficient for repression. The KAP-1 repression domains comprised of the HP1 binding domain, the PHD domain and the bromo domain, which together function to assemble the machinery required for repression. A major goal has been to define this machinery "downstream" of KAP-1, which is required and, in light of the evidence presented here, whether KAP-1 oligomerization also plays a role in this recruitment. A likely direct target of KAP-1 in this context was recently defined by us and others as the heterochromatin protein HP1 (Ryan *et al.*, 1999). The HP1 proteins are a small family of low molecular mass, non-histone chromosomal proteins, some of which are tightly associated with silenced heterochromatin. We have shown that KAP-1 directly binds to purified recombinant HP1 protein, and that a stable quaternary complex can be formed between DNA, a KRAB-ZFP, KAP-1 and HP1. Mutations in KAP-1 that abolish HP1 binding significantly impair KAP-1-mediated repression. Furthermore, HP1 and KAP-1 extensively colocalize to heterochromatic regions in interphase nuclei. Thus, it is tempting to speculate that KAP-1 serves to nucleate HP1-mediated stable repression on a template, and that the oligomerization of KAP-1 as it binds the KRAB domain has a role in the nucleation of heterochromatin.

In summary, we have provided the first analysis of a direct protein-protein interaction involving an RBCC protein and its physiologically relevant protein ligand. These studies have highlighted the apparently high specificity for recognition inherent in this protein-protein interface. Further refinement of these macromolecular interactions could lead to strategies for harnessing KAP-1-mediated repression *in vivo* using synthetic transcription factors for possible therapeutic benefit. One could also envision a small-molecule inhibitor strategy directed at the KRAB-KAP-1 system, which may be useful for reprogramming the expression of stably repressed genes and thus altering cellular phenotypes.

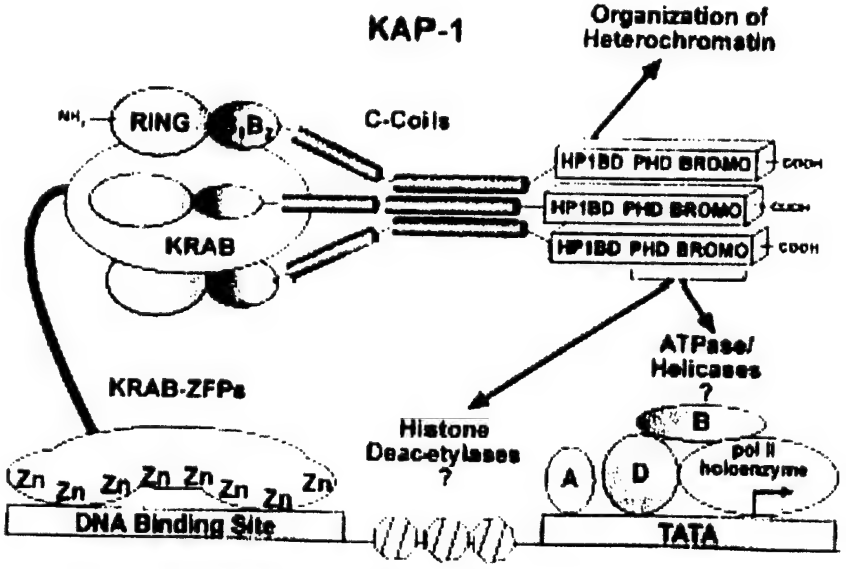



**Figure 10.** Model for the KRAB-KAP-1 interaction. This model is described in Discussion.

## Materials and Methods

### Preparation of plasmids

The plasmid expressing the 6His GAL4-KRAB (1-90) protein was constructed *via* polymerase chain reaction (PCR) using the pM1-KOX, 1-90 plasmid (Margolin *et al.*, 1994) as a template. A 5′ oligonucleotide, which incorporated a *Bam*HI site immediately 5′ to the Gal4 initiator methionine residue and a 3′ oligonucleotide, which incorporated a stop codon and a *Hin*dIII site after amino acid residue 90 of KOX-1, were used to amplify the desired sequence. The digested PCR product was cloned into the corresponding restriction sites of the pQE30 vector (Qiagen).

The plasmid expressing the 6His KRAB (1-90) protein was produced by subcloning the DNA segment encoding residues 1-90 of KOX1-KRAB from pM1 into the pSP73 vector, using the 5′ *Eco*RI and 3′ *Xba*I sites. This fragment was then excised and subcloned *via* the flanking *Bam*HI and *Pst*I sites in pSP73 into the corresponding sites of the pQE30 vector. The protein expressed from this plasmid contained the 18 $NH_2$-terminal amino acid residues, MRGSHHHHHHGSDIIDEF, followed by residues 1-90 of the KOX1 protein, then the C-terminal sequence RVDLQPSLIS is present.

The pQE30 KAP-1-RBCC plasmid was constructed by amplifying the DNA fragment encoding residues 22-418 of human KAP-1 using the primer pair, forward 5′ CGT TGA ATT CTA CCA TGG CTT CGT GGG GAT CCC 3′, reverse 5′ CGT AAA GCT TGC CTG TTG AGT TAG TGC C 3′. The PCR product was cleaved with *Eco*RI and *Hin*dIII, and cloned into pBluescript II KS +/− to generate pBS-KS KAP-1-RBCC. This fragment was then excised with the flanking 5′ *Sma*I site and 3′ *Hin*dIII sites in pBluescript II KS +/− and cloned into pQE30 plasmid. The expressed KAP-1-RBCC protein contained the 18 $NH_2$-terminal amino acid residues, MRGSHHHHHHIGSACELGT, followed by residues 22-418 of the KAP-1 protein, then the C-terminal sequence KLN. The point mutations within the RBCC domain of KAP-1 were generated in this plasmid using PCR-mediated mutagenesis. The mutagenic primers contained the following codons: (CC65,68AA), TGC to GCC; (CCC152,153,156AAA), TGC to GCC, TGT to GCT; (C209A,H212A), TGC to GCC, CAC to GCC; (L270P), TTG to CCG; (L306P), CTG to CCG; (V260N), GTG to AAC; (V260D), GTG to GAC.

A baculovirus vector (pVL1392) encoding residues 22-418 of KAP-1 was generated by first digesting pBS-KS KAP-1-RBCC with *Hin*dIII. The overhang in the linearized plasmid was filled in with Klenow enzyme, and then the DNA was digested with *Eco*RI. This fragment was cloned into pVL1392 (PharMingen) at the *Eco*RI and *Sma*I sites. The protein expressed contained the 16 $NH_2$-terminal amino acid residues, MASWGSHHHHHHDPGY, followed by residues 22-418 of the KAP-1 protein, then the C-terminal sequence KLGSFPGTRQEPKTHSLQGNP.

The pQE30 TIF1α-RBCC plasmid was produced by amplifying the DNA fragment encoding residues 1-414 of TIF1α using the primer pair; forward 5′ GCA TAG ATC TAT GGA GGT GGC TGT GGA GAA G 3′; reverse 5′ CGA TGC ATG CTC AGA GCC CAG AAA CTA GGA TC 3′. The PCR product was digested with *Bgl*II and *Sph*I, and cloned into the *Bam*HI and *Sph*I sites of the pQE30 vector.

The plasmids expressing the KAP-1 RB region fused to various heterologous oligomerization domains and KAP-1/MID-1 domain swaps were generated by overlap-extension PCR-mediated mutagenesis. The GCN4-LZ and GCN4-TZ domains have been described and spanned residues 253-281 and 250-281 of GCN4, respectively (Waterman *et al.*, 1996). The *Saccharomyces cerevisiae* heat-shock transcription factor (HSF) trimerization domain spanning residues 321-441 was used (Drees *et al.*, 1997). The forward primers are: KAP-RB-[$HSF_{Tri}$], KAP-RB-[$GCN4_{Dim}$], KAP-RB-[$GCN4_{Tet}$], $RB_{KAP}$-$CC_{MID}$, $R_{KAP}$-$BCC_{MID}$, 5′ AGC GGA TTC GAA TTC ACC ACC ATG GCT CAC CAT CAC CAC CAT CAC 3′ (*Bam*HI); MID-RBCC, $R_{MID}$-$BCC_{KAP}$, 5′ GCA TAG ATC TAT GGA AAC ACT GGA GTC AG 3′ (*Bgl*II). The reverse primers are: KAP-RB-[$HSF_{Tri}$], 5′ CGT TAA GCT TTC AAG TGC CTG TGT TGT CGC G 3′ (*Hin*dIII); KAP-RB-[$GCN4_{Dim}$], 5′ CGT TAA GCT TTC AGC GTT CGC CAA CTA A 3′ (*Hin*dIII); KAP-RB-[$GCN4_{Tet}$], 5′ CGT TAA GCT TTC AAC GTT CAC CCA ATA ATT T 3′ (*Hin*dIII); MID-RBCC, 5′ CGA TGC ATG CTC AAT CTA AGG CAA AGG TGT C 3′ (*Sph*I); $RB_{KAP}$-$CC_{MID}$, $R_{KAP}$-$BCC_{MID}$, 5′ GAT CAA GCT TTC AAT CTA AGG CAA AGG TGT C 3′ (*Hin*dIII); $R_{MID}$-$BCC_{KAP}$, 5′ ATG CAA GCT TTC ATG TTG AGT TAG TGC CAG GAC 3′ (*Hin*dIII). Fusion primers (forward) are: KAP-RB-[$HSF_{Tri}$], 5′ TGA GGA ACC AGC GCA AGA ATG GTC ATT TAT TGC AGG 3′; KAP-RB-[$GCN4_{Dim}$], 5′ TGA GGA ACC AGC GCA AGA TCC GCG GGC GTC AGC TT 3′; KAP-RB-[$GCN4_{Tet}$], 5′ TGA GGA ACC AGC GCA AGA TCC GCG GTA ATC GTC TG 3′; $RB_{KAP}$-$CC_{MID}$, 5′ GTG AGG AAC CAG CGC AAG AAG CAA AAC TTA GAG AGT AAC 3′; $R_{KAP}$-$BCC_{MID}$, 5′ AAT TAT TTC ATG CGT GAT AGT GGG CTC AAG CGC AAC GTC ACC 3′; $R_{MID}$-$BCC_{KAP}$, 5′ CTC AGC CAG CGA GGT CTA GAC GGC AGC AAG GCT GCC ACC GAC 3′. The second-round PCR products were digested with the appropriate enzymes defined for each fusion in parentheses above and subcloned into compatible restriction sites of the pQE30 vector. All plasmids generated by PCR were confirmed by DNA sequencing on both strands to verify the presence of desired mutations and appropriate fusions (Wistar DNA Core Facility).

### Protein purification

*Escherichia coli* SG13009 cells (Qiagen) bearing the desired plasmid were propagated with aeration at 37°C in one liter of Luria broth to an $A_{600}$ of approximately 0.6. IPTG was added to 1 mM, and growth at 37°C was continued for four hours. The cells were harvested by centrifugation.

The purification of KOX1-KRAB protein was performed at room temperature (Shi *et al.*, 1997). For one liter of cell culture, the cell pellet was resuspended in 40 ml of P300 buffer (10 mM $Na_2HPO_4$, 1.4 mM $KH_2PO_4$, 2.7 mM KCl, 450 mM NaCl, pH 7.0, 1 mM phenylmethylsulfonyl fluoride (PMSF)). The cells were lysed by sonication on ice and the cell extract was centrifuged at 12,000 *g* for 20 minutes. The pellet was dissolved in 40 ml of lysis buffer (20 mM Tris-HCl (pH 7.8), 0.5 M NaCl, 5 mM imidazole, 8 M urea), and incubated for one hour. The lysate was then centrifuged at 12,000 *g* for 20 minutes at 4°C. To the supernatant, 1-1.5 ml packed column volume of Ni-NTA resin (Qiagen), previously equilibrated in lysis buffer, was added in batch and the incubation continued with rotation for one hour. The resin was then washed with lysis buffer and loaded onto a disposable column, followed by an additional wash buffer (20 mM Tris-HCl (pH 7.8), 0.5 M NaCl, 20 mM imidazole, 8 M urea). The KRAB protein bound to the

nickel column was renatured on the column by stepping the urea in lysis buffer from 8 M to 0 M. The renatured KRAB protein was eluted with elution buffer (20 mM Tris-HCl (pH 7.8), 0.5 M NaCl, 0.3 M imidazole). The protein was dialyzed against 20 mM Tris-HCl (pH 8.0), 0.5 M NaCl, followed by a second dialysis reducing the salt to 50 mM NaCl.

The KAP-1-RBCC proteins, mutants thereof, and the chimeras, were purified using both denaturing and non-denaturing conditions. The purification was initially performed at 4 C under non-denaturing condition as recommended by the manufacturer (Qiagen). The bacterial pellet was resuspended in P300 buffer (pH 7.0), 10% glycerol, 10 mM imidazole, and lysed by sonication. The cell extract was centrifuged at 12,000 *g* for 20 minutes. The insoluble pellet was then used for purification under denaturing conditions (see below). The supernatant fraction containing soluble protein was incubated in batch with Ni-NTA resin for one hour. The resin was washed with P300 buffer. Then the resin was loaded into a disposable column, and washed with P300 buffer at pH 6.0. The protein was eluted with P300 (pH 7.0), 10% (v/v) glycerol, 300 mM imidazole. The eluted protein was dialyzed against three changes of P300 (pH 7.0), 10% glycerol, 20 μM $ZnSO_4$, 0.5 mM DTT).

The insoluble proteins from the pellet fraction were purified and eluted from the Ni-NTA with 300 mM imidazole and low pH under denaturing condition as recommended by the manufacturer. The eluted proteins were diluted to 20 ml with P300, and dialyzed against P300 (pH 7.0), 10% glycerol, 20 μM $ZnSO_4$, 0.5 mM DTT. After dialysis, the protein was clarified by centrifugation and concentrated in an Amicon concentrator (Amicon) following the manufacturer's instruction.

### Electrophoretic mobility shift assay (EMSA)

EMSA assays were performed essentially as described (Ryan *et al.*, 1999; Fredericks *et al.*, 1995). The purified recombinant GAL4-KRAB or GAL4-KRAB (DV-AA) protein was incubated in binding buffer with either purified bacterial KAP-1-RBCC, KAP-1-RBCC mutant/fusion proteins, or baculovirus whole-cell lysates for 15 minutes at 30 C, and then 1 μl of $^{32}$P-labeled Gal4 probe ( $10^5$ cpm/μl) was added and the reaction was incubated for an additional 15 minutes at 30 C. The DNA:protein complexes were resolved on native polyacrylamide gels by electrophoresis at 450 V for one hour and 45 minutes in 45 mM Tris-borate (pH 8.3), 1 mM EDTA buffer at 4 C. The EMSA gels were dried and visualized by autoradiography. The Gal4 was the double-stranded synthetic oligonucleotide 5' GAT CCC GGA GGA CAG TAC TCC GT 3', which was labeled with [$^{32}$P]ATP as described (Fredericks *et al.*, 1995).

### Reconstitution of the KOX1-KRAB:KAP-1-RBCC complex *in vitro*

To reconstitute the KOX1-KRAB:KAP-1-RBCC complex *in vitro*, highly purified bacterial KOX1-KRAB and KAP-1-RBCC proteins were incubated at a 1:1 molar ratio in buffer containing 20 mM Hepes, 300 mM NaCl, 0.5 mM DTT, 10% glycerol at 30°C for 30 minutes. The samples were chilled at 4°C for five minutes, and then centrifuged at 10,000 *g* for 5 minutes at 4°C. The supernatant was analyzed by gel-filtration, sucrose gradient sedimentation, and chemical cross-linking.

### Gel-filtration analysis

Gel-filtration of highly purified proteins was performed using a Superdex 200 HR 10/30 column equilibrated in P300 buffer (pH 7.0), 10% glycerol, 0.1% NP-40, 20 μM $ZnSO_4$, 1 mM DTT, using an FPLC system (Pharmacia). The column was run at 4 C at a flow-rate of 0.5 ml/minute. and 1 ml fractions were collected. Aliquots of each fraction were analyzed by SDS-PAGE or Western blot (Ryan *et al.* 1999).

### Sucrose gradient sedimentation

Purified bacterial KAP-1-RBCC protein (60 μg), KOX1-KRAB (60 μg) or, KOX1-KRAB (60 μg):KAP-1-RBCC (200 μg) complex in 200 μl of sucrose gradient buffer (10 mM sodium phosphate (pH 7.0), 140 mM NaCl, 0.5% NP-40, 1 mM PMSF, 2 μg/ml aprotinin, 3 μg/ml leupeptin) was layered onto a discontinuous 5% to 60% (w/v) sucrose gradient. The gradients were pre-formed at room temperature for one hour, followed by an additional hour at 4 C prior to loading the sample. The sucrose gradients were centrifuged for 16 hours at 4 C in an SW-50.1 rotor at 190,000 *g*. Approximately 20 fractions (250 μl each) were collected from the base of each centrifuge tube. Aliquots of each fraction were subjected to SDS-PAGE, and the proteins were stained with Coomassie brilliant blue.

### Non-denaturing (ND)-PAGE

ND-PAGE was conducted by the procedure described in the ND protein Molecular Weight Determination Kit from Sigma. Purified bacterial KAP-1-RBCC protein (10 μg) and 10 μg of each protein standard were analyzed by ND-PAGE on gels of various polyacrylamide concentrations (6, 7, 8, and 10%), and the proteins were stained with Coomassie brilliant blue.

### Analytical ultracentrifugation

The KAP-1-RBCC (0.7 to 2.0 mg/ml) sample was prepared for analytical ultracentrifugation by dialysis against a P300 buffer (pH 7.0), 10% (v/v) glycerol, 20 μM $ZnSO_4$, 1 mM DTT. The sedimentation equilibrium experiments were performed in an Optima XL-I ultracentrifuge, using the interference optics to determine the protein concentration gradient. For each experiment, three cells were assembled with double-sector 12 mm centerpieces and sapphire windows, and loaded with 110 μl of reference buffer (dialysate) and 110 μl of sample, at three different protein concentrations. A blank scan of distilled water was taken before the run, to correct for the effects of window distortion on the fringe displacement data (Yphantis, 1964). The experiments were performed at 4-25°C, at a speed of 13,000 rpm. Fringe displacement data were collected every six to eight hours until equilibrium was reached, as determined by comparison of successive scans using the MATCH program, and the data were edited using the REEDIT program (both programs kindly provided by David Yphantis).

Analysis of sedimentation equilibrium data was performed using the NONLIN program (Johnson *et al.*, 1981). The partial specific volume of KAP-1-RBCC was calculated from the amino acid sequence, and the density of the solvent was calculated according to Laue *et al.* (1992). Three data sets from different loading concen-

trations were fitted simultaneously. Goodness of fit was determined by examination of the residuals and minimization of the variance.

### Chemical cross-linking

Chemical cross-linking of recombinant proteins using ethylene glycol bis (succinimidylsuccinate) (EGS) (Pierce) was performed essentially as described (Drees *et al.*, 1997) with the following modifications. For analysis of the oligomerization of wild-type or mutant KAP-1-RBCC proteins, or KOX1-KRAB:KAP-1-RBCC complexes, the final protein concentration was 5 μM in each reaction. The EGS cross-linking was performed at room temperature for five minutes in P300 buffer (pH 7.0), 10% glycerol, 1 mM DTT. The EGS was freshly prepared in dimethylsulfoxide and was added from a stock solution to final concentrations ranging from 0.078 to 5 mM. The reactions were terminated by addition of lysine to a final concentration of 20 mM. After cross-linking, 10 μl of 5× Laemmli sample buffer was added to the reactions, and the samples were analyzed by SDS-PAGE. The proteins was detected by Western blot analysis as described (Ryan *et al.* 1999).

### Expression of KAP-1 and KAP-1-RBCC in baculovirus-infected cells

The KAP-1-RBCC proteins and the full-length KAP-1 protein were expressed individually or co-expressed in baculovirus-infected Sf9 cells (Wistar Core Protein Expression Facility). High-titer baculovirus stocks for each protein were prepared using standard technique and Sf9 cells were infected for 48 hours. For a 50 ml cell culture, the cell pellets were lysed in 5 ml of 10 mM Tris-HCl (pH 8.0), 150 mM NaCl, 1% NP-40, 1 mM benzamidine, 1 mM PMSF, 2 mM leupeptin, 10 μg/ml pepstatin, 10 μg/ml aprotinin, and 10 mM NaF. The cells were lysed at 4°C for 20 minutes, followed by centrifugation at 16,000 *g* for ten minutes. The protein was then quantified by Bradford assay (Bio Rad). For protein-protein association assays, the supernatant fraction (250 μg protein) of an individual protein preparation was added to 50 μl of 50% (w/v) Ni-NTA resin slurry, and incubated at 4°C for one hour. The resin was washed with P300 (pH 7.0), 0.1% NP-40, 1 mM PMSF. The bound proteins were extracted from resin with 5 × Laemmli sample buffer, resolved by SDS-PAGE, and detected by Western blot (Ryan *et al.* 1999).

## Acknowledgments

H.P. and D.E.J. were supported by the Wistar Basic Cancer Research Training Grant CA 09171. D.C.S. is supported by DAMD17-98-1-8269. D.W.S. is supported by CA74294 and CA66671. F.J.R. is supported, in part, by National Institutes of Health grants CA 52009, Core grant CA 10815, DK 49210, GM 54220, DAMD 17-96-1-6141, ACS NP-954, the Irving A. Hansen Memorial Foundation, the Mary A. Rumsey Memorial Foundation, and the Pew Scholars Program in the Biomedical Sciences. We thank H.C.M. Nelson for the HSF trimerization domain plasmid, and T.D. Halazonetis for the GCN4-LZ and GCN4-TZ domains plasmid. We thank G. Meroni and A. Ballabio for the MID1 plasmid, and M. Lazar for the TIF1α plasmid. We thank R. Ryan for constructing the bv.KAP-1 plasmid, and members of the Speicher, Marmorstein, Berger, and Halazonetis laboratories for many helpful discussions.

## References

Barlow, P. N., Luisi, B., Milner, A., Elliott, M. & Everett, R. (1994). Structure of the C3HC4 domain by 1H-nuclear magnetic resonance spectroscopy. A new structural class of zinc-finger. *J. Mol. Biol.* **237**, 201-211.

Baxevanis, A. D. & Vinson, C. R. (1993). Interactions of coiled coils in transcription factors: where is the specificity? *Curr. Opin. Genet. Dev.* **3**, 278-285.

Bellefroid, E. J., Poncelet, D. A., Lecocq, P. J., Revelant, O. & Martial, J. A. (1991). The evolutionarily conserved Kruppel-associated box domain defines a subfamily of eukaryotic multifingered proteins. *Proc. Natl Acad. Sci. USA*, **88**, 3608-3612.

Bellini, M., Lacroix, J. C. & Gall, J. G. (1995). A zinc-binding domain is required for targeting the maternal nuclear protein PwA33 to lampbrush chromosome loops. *J. Cell. Biol.* **131**, 563-570.

Bellon, S. F., Rodgers, K. K., Schatz, D. G., Coleman, J. E. & Steitz, T. A. (1997). Crystal structure of the RAG1 dimerization domain reveals multiple zinc-binding motifs including a novel zinc binuclear cluster. *Nature Struct. Biol.* **4**, 586-591.

Boddy, M. N., Duprez, E., Borden, K. L. & Freemont, P. S. (1997). Surface residue mutations of the PML RING finger domain alter the formation of nuclear matrix-associated PML bodies. *J. Cell. Sci.* **110**, 2197-2205.

Borden, K. L. (1998). RING fingers and B-boxes: zinc-binding protein-protein interaction domains. *Biochem. Cell. Biol.* **76**, 351-358.

Borden, K. L., Boddy, M. N., Lally, J., O'Reilly, N. J., Martin, S., Howe, K., Solomon, E. & Freemont, P. S. (1995a). The solution structure of the RING finger domain from the acute promyelocytic leukaemia proto-oncoprotein PML. *EMBO J.* **14**, 1532-1541.

Borden, K. L., Lally, J. M., Martin, S. R., O'Reilly, N. J., Etkin, L. D. & Freemont, P. S. (1995b). Novel topology of a zinc-binding domain from a protein involved in regulating early Xenopus development. *EMBO J.* **14**, 5947-5956.

Borden, K. L., Lally, J. M., Martin, S. R., O'Reilly, N. J., Solomon, E. & Freemont, P. S. (1996). In vivo and in vitro characterization of the B1 and B2 zinc-binding domains from the acute promyelocytic leukemia protooncoprotein PML. *Proc. Natl Acad. Sci. USA*, **93**, 1601-1606.

Borden, K. L., Campbell-Dwyer, E. J. & Salvato, M. S. (1997). The promyelocytic leukemia protein PML has a pro-apoptotic activity mediated through its RING domain. *FEBS Letters*, **418**, 30-34.

Bryan, J. K. (1977). Molecular weights of protein multimers from polyacrylamide gel electrophoresis. *Anal. Biochem.* **78**, 513-519.

Buchner, G., Montini, E., Andolfi, G., Quaderi, N., Cainarca, S., Messali, S., Bassi, M. T., Ballabio, A., Meroni, G. & Franco, B. (1999). MID2, a homologue of the opitz syndrome gene MID1: similarities in subcellular localization and differences in expression during development. *Hum. Mol. Genet.* **8**, 1397-1407.

Cainarca, S., Messali, S., Ballabio, A. & Meroni, G. (1999). Functional characterization of the opitz syn-

drome gene product (midin): evidence for homodimerization and association with microtubules throughout the cell cycle. *Hum. Mol. Genet.* **8**, 1387-1396.

Cao, T., Borden, K. L. B., Freemont, P. S. & Etkin, L. D. (1997). Involment of the rfp tripartite motif in protein-protein interactions and subcellular distribution. *J. Cell Sci.* **110**, 1563-1571.

Cao, T., Duprez, E., Borden, K. L., Freemont, P. S. & Etkin, L. D. (1998). Ret finger protein is a normal component of PML nuclear bodies and interacts directly with PML. *J. Cell Sci.* **111**, 1319-1329.

Drees, B. L., Grotkopp, E. K. & Nelson, H. C. (1997). The GCN4 leucine zipper can functionally substitute for the heat shock transcription factor's trimerization domain. *J. Mol. Biol.* **273**, 61-74.

Dyck, J. A., Maul, G. G., Miller, W. H., Jr, Chen, J. D., Kakizuka, A. & Evans, R. M. (1994). A novel macromolecular structure is a target of the promyelocyte-retinoic acid receptor oncoprotein. *Cell*, **76**, 333-343.

El-Husseini, A. E. & Vincent, S. R. (1999). Cloning and characterization of a novel RING finger protein that interacts with class V myosins. *J. Biol. Chem.* **274**, 19771-19777.

Fagioli, M., Alcalay, M., Tomassoni, L., Ferrucci, P. F., Mencarelli, A., Riganelli, D., Grignani, F., Pozzan, T., Nicoletti, I. & Pelicci, P. G. (1998). Cooperation between the RING + B1-B2 and coiled-coil domains of PML is necessary for its effects on cell survival. *Oncogene*, **16**, 2905-2913.

Fredericks, W. J., Galili, N., Mukhopadhyay, S., Rovera, G., Bennicelli, J., Barr, F. G. & Rauscher, F., Jr (1995). The PAX3-FKHR fusion protein created by the t(2;13) translocation in alveolar rhabdomyosarcomas is a more potent transcriptional activator than PAX3. *Mol. Cell. Biol.* **15**, 1522-1535.

Friedman, J. R., Fredericks, W. J., Jensen, D. E., Speicher, D. W., Huang, X. P., Neilson, E. G. & Rauscher, F., Jr (1996). KAP-1, a novel corepressor for the highly conserved KRAB repression domain. *Genes Dev.* **10**, 2067-2078.

Goddard, A. D., Borrow, J., Freemont, P. S. & Solomon, E. (1991). Characterization of a zinc finger gene disrupted by the t(15;17) in acute promyelocytic leukemia. *Science*, **254**, 1371-1374.

Grignani, F., Testa, U., Rogaia, D., Ferrucci, P. F., Samoggia, P., Pinto, A., Aldinucci, D., Gelmetti, V., Fagioli, M., Alcalay, M., Seeler, J., Nicoletti, I., Peschle, C. & Pelicci, P. G. (1996). Effects on differentiation by the promyelocytic leukemia PML/RARalpha protein depend on the fusion of the PML protein dimerization and RARalpha DNA binding domains. *EMBO J.* **15**, 4949-4958.

Hasegawa, N., Iwashita, T., Asai, N., Murakami, H., Iwata, Y., Isomura, T., Goto, H., Hayakawa, T. & Takahashi, M. (1996). A RING finger motif regulates transforming activity of the rfp/ret fusion gene. *Biochem. Biophys. Res. Commun.* **225**, 627-631.

Hemenway, C. S., Halligan, B. W. & Levy, L. S. (1998). The Bmi-1 oncoprotein interacts with dinG and MPh2: the role of RING finger domains. *Oncogene*, **16**, 2541-2547.

Hope, I. A. & Struhl, K. (1987). GCN4, a eukaryotic transcriptional activator protein, binds as a dimer to target DNA. *EMBO J.* **6**, 2781-2784.

Hurst, H. C. (1995). Transcription factors 1: bZIP proteins. *Protein Profile*, **2**, 101-168.

Irminger-Finger, I., Siegel, B. D. & Leung, W. C. (1999). The functions of breast cancer susceptibility gene 1 (BRCA1) product and its associated proteins. *Biol. Chem.* **380**, 117-128.

Johnson, M. L., Correia, J. J., Yphantis, D. A. & Halvorson, H. R. (1981). Analysis of data from the analytical ultracentrifuge by nonlinear least- squares techniques. *Biophys. J.* **36**, 575-588.

Kakizuka, A., Miller, W. H., Jr, Umesono, K., Warrell, R. P., Jr, Frankel, S. R., Murty, V. V., Dmitrovsky, E. & Evans, R. M. (1991). Chromosomal translocation t(15;17) in human acute promyelocytic leukemia fuses RAR alpha with a novel putative transcription factor, PML. *Cell*, **66**, 663-674.

Laue, T. M., Shah, B., Ridgeway, T. M. & Pelletier, S. L. (1992). Computer-aided interpretation of analytical sedimentation data for proteins. In *Analytical Ultracentrifugation in Biochemistry and Polymer Science* (Harding, S. E., Rowe, A. J. & Horton, J. C., eds), pp. 90-125, Royal Society of Chemistry, Cambridge.

Le, Douarin B., Zechel, C., Garnier, J. M., Lutz, Y., Tora, L., Pierrat, P., Heery, D., Gronemeyer, H., Chambon, P. & Losson, R. (1995). The N-terminal part of TIF1, a putative mediator of the ligand-dependent activation function (AF-2) of nuclear receptors, is fused to B-raf in the oncogenic protein T18. *EMBO J.* **14**, 2020-2033.

Lupas, A. (1996). Coiled coils: new structures and new functions. *Trends Biochem. Sci.* **21**, 375-382.

Margolin, J. F., Friedman, J. R., Meyer, W. K., Vissing, H., Theisen, H. J. & Rauscher, F., Jr (1994). Kruppel-associated boxes are potent transcriptional repression domains. *Proc. Natl Acad. Sci. USA*, **91**, 4509-4513.

Maul, G. G. (1998). Nuclear domain 10, the site of DNA virus transcription and replication. *Bioessays*, **20**, 660-667.

McWhirter, S. M., Pullen, S. S., Holton, J. M., Crute, J. J., Kehry, M. R. & Alber, T. (1999). Crystallographic analysis of CD40 recognition and signaling by human TRAF2. *Proc. Natl Acad. Sci. USA*, **96**, 8408-8413.

Meza, J. E., Brzovic, P. S., King, M. C. & Klevit, R. E. (1999). Mapping the functional domains of BRCA1. Interaction of the ring finger domains of BRCA1 and BARD1. *J. Biol. Chem.* **274**, 5659-5665.

Miki, T., Fleming, T. P., Crescenzi, M., Molloy, C. J., Blam, S. B., Reynolds, S. H. & Aaronson, S. A. (1991). Development of a highly efficient expression cDNA cloning system: application to oncogene isolation. *Proc. Natl Acad. Sci. USA*, **88**, 5167-5171.

Moosmann, P., Georgiev, O., Le, Douarin B., Bourquin, J. P. & Schaffner, W. (1996). Transcriptional repression by RING finger protein TIF1 beta that interacts with the KRAB repressor domain of KOX1. *Nucl. Acids Res.* **24**, 4859-4867.

Ogawa, S., Goto, W., Orimo, A., Hosoi, T., Ouchi, Y., Muramatsu, M. & Inoue, S. (1998). Molecular cloning of a novel RING finger-B box-coiled coil (RBCC) protein, terf, expressed in the testis. *Biochem. Biophys. Res. Commun.* **251**, 515-519.

Park, Y. C., Burkitt, V., Villa, A. R., Tong, L. & Wu, H. (1999). Structural basis for self-association and receptor recognition of human TRAF2. *Nature*, **398**, 533-538.

Pengue, G., Caputo, A., Rossi, C., Barbanti-Brodano, G. & Lania, L. (1995). Transcriptional silencing of human immunodeficiency virus type 1 long terminal repeat-driven gene expression by the Kruppel-associated box repressor domain targeted to the

transactivating response element. *J. Virol.* **69**, 6577-6580.

Perez, A., Kastner, P., Sethi, S., Lutz, Y., Reibel, C. & Chambon, P. (1993). PMLRAR homodimers: distinct DNA binding properties and heteromeric interactions with RXR. *EMBO J.* **12**, 3171-3182.

Perrin, K. & Lacroix, J. C. (1998). XL43 and XL75: two novel RING finger-containing genes expressed during oogenesis and embryogenesis in Xenopus laevis. *Gene*, **210**, 127-134.

Pullen, S. S., Labadia, M. E., Ingraham, R. H., McWhirter, S. M., Everdeen, D. S., Alber, T., Crute, J. J. & Kehry, M. R. (1999). High-affinity interactions of tumor necrosis factor receptor-associated factors (TRAFs) and CD40 require TRAF trimerization and CD40 multimerization. *Biochemistry*, **38**, 10168-10177.

Quaderi, N. A., Schweiger, S., Gaudenz, K., Franco, B., Rugarli, E. I., Berger, W., Feldman, G. J., Volta, M., Andolfi, G., Gilgenkrantz, S., Marion, R. W., Hennekam, R. C., Opitz, J. M., Muenke, M., Ropers, H. H. & Ballabio, A. (1997). Opitz G/BBB syndrome, a defect of midline development, is due to mutations in a new RING finger gene on Xp22. *Nature Genet.* **17**, 285-291.

Reddy, B. A. & Etkin, L. D. (1991). A unique bipartite cysteine-histidine motif defines a subfamily of potential zinc-finger proteins. *Nucl. Acids Res.* **19**, 6330.

Ryan, R. F., Schultz, D. C., Ayyanathan, K., Singh, P. B., Friedman, J. R., Fredericks, W. J. & Rauscher, F. J., III (1999). KAP-1 corepressor protein interacts and colocalizes with heterochromatic and euchromatic HP1 proteins: a potential role for Kruppel-associated box-zinc finger proteins in heterochromatin-mediated gene silencing. *Mol. Cell. Biol.* **19**, 4366-4378.

Saurin, A. J., Borden, K. L., Boddy, M. N. & Freemont, P. S. (1996). Does this have a familiar RING? *Trends Biochem. Sci.* **21**, 208-214.

Shi, P. Y., Maizels, N. & Weiner, A. M. (1997). Recovery of soluble active recombinant protein from inclusion bodies. *Biotechniques*, **23**, 1036-1038.

Sternsdorf, T., Grotzinger, T., Jensen, K. & Will, H. (1997). Nuclear dots: actors on many stages. *Immunobiology*, **198**, 307-331.

Takahashi, M., Ritz, J. & Cooper, G. M. (1985). Activation of a novel human transforming gene, ret, by DNA rearrangement. *Cell*, **42**, 581-588.

Tatematsu, K., Tokunaga, C., Nakagawa, N., Tanizawa, K., Kuroda, S. & Kikkawa, U. (1998). Transcriptional activity of RBCK1 protein (RBCC protein interacting with PKC 1): requirement of RING-finger and B-Box motifs and regulation by protein kinases. *Biochem. Biophys. Res. Commun.* **247**, 392-396.

Venturini, L., You, J., Stadler, M., Galien, R., Lallemand, V., Koken, M. H., Mattei, M. G., Ganser, A., Chambon, P., Losson, R. & de The, H. (1999). TIF1-gamma, a novel member of the transcriptional intermediary factor 1 family. *Oncogene*, **18**, 1209-1217.

Vissing, H., Meyer, W. K., Aagaard, L., Tommerup, N. & Thiesen, H. J. (1995). Repression of transcriptional activity by heterologous KRAB domains present in zinc finger proteins. *FEBS Letters*, **369**, 153-157.

Waterman, M. J., Waterman, J. L. & Halazonetis, T. D. (1996). An engineered four-stranded coiled coil substitutes for the tetramerization domain of wild-type p53 and alleviates transdominant inhibition by tumor-derived p53 mutants. *Cancer Res.* **56**, 158-163.

Weis, K., Rambaud, S., Lavau, C., Jansen, J., Carvalho, T., Carmo-Fonseca, M., Lamond, A. & Dejean, A. (1994). Retinoic acid regulates aberrant nuclear localization of PML-RAR alpha in acute promyelocytic leukemia cells. *Cell*, **76**, 345-356.

Witzgall, R., O'Leary, E., Leaf, A., Onaldi, D. & Bonventre, J. V. (1994). The Kruppel-associated box-A (KRAB-A) domain of zinc finger proteins mediates transcriptional repression. *Proc. Natl Acad. Sci. USA*, **91**, 4514-4518.

Wolf, E., Kim, P. S. & Berger, B. (1997). MultiCoil: a program for predicting two- and three-stranded coiled coils. *Protein Sci.* **6**, 1179-1189.

Yphantis, D. A. (1964). Equilibrium ultracentrifugation of dilute solutions. *Biochemistry*, **3**, 297-317.

***Edited by P. E. Wright***

*(Received 22 September 1999; received in revised form 17 November 1999; accepted 18 November 1999)*

# A POTENTIAL ROLE FOR THE KRAB-ZFP:KAP-1 TRANSCRIPTIONAL REPRESSION COMPLEX IN HETEROCHROMATIN MEDIATED GENE SILENCING.

David C. Schultz, Mark S. Lechner, Hongzhuang Peng, Kasirajan Ayyanathan, Geeta Patel, and Frank J. Rauscher, III. The Wistar Institute, Philadelphia, PA 19104

Nearly one-third of all C2H2 zinc finger proteins possess the highly conserved KRAB (Kruppel-associated box) domain which functions as a potent transferable, DNA-binding dependent transcriptional repression domain. We have previously cloned a universal corepressor, KAP-1, which is required for KRAB-mediated repression *in vivo*. Structurally, KAP-1 is the proto-type for an emerging superfamily of corepressors, consisting of a RING finger, two B boxes, and a coiled-coil domain which collectively are necessary and sufficient for the interaction with the KRAB domain. Transcriptional repression by the KRAB:KAP-1 complex is independent of histone deacetylase, as no stably associated deacetylase activity or interaction with well characterized components of histone deacetylase complexes has been observed. We have dissected the repression activity of the KAP-1 corepressor into at least two non-overlapping repression domains. The PHD and bromodomain together display DNA-binding dependent repression that can be disrupted by specific amino acid substitutions in either motif. A separate repression domain within KAP-1 is composed of 38 amino acids which binds directly to the chromo-shadow domains of both human and murine heterochromatin proteins (HP1). Specific amino acid substitutions within this HP1 binding domain can abolish HP1 binding *in vitro* and reduce the repression activity of KAP-1 *in vivo*. Immuno-localization studies in interphase nuclei of NIH/3T3 cells revealed that KAP-1 staining is almost exclusively nuclear, excluded from nucleoli, and concentrated in islands of pericentromeric heterochromatin. Furthermore, this KAP-1 staining pattern is quite dynamic during the cell cycle with tight colocalization with pericentromeric heterochromatin during interphase which diminishes as the cell enters mitosis. Dual staining immunofluoresence studies revealed that KAP-1 colocalized with human and murine HP1 proteins in interphase nuclei, lending additional support to the biochemical data that HP1's interact with KAP-1 *in vivo*. These data suggest a repression mechanism by which the KRAB-ZFP/KAP-1 complex may recruit HP1 proteins to specific loci within the genome to form heterochromatin-like complexes that silence gene activity. We speculate that gene-specific repression may be a consequence of the formation of such complexes, ultimately leading to silenced genes in newly formed heterochromatic chromosomal environments.

**#858 Characterization of the KAP-1 corepressor interaction with HP1 proteins.** Lechner, M.S., Schultz, D.C., Ayyanathan, K., *Singh, P., and Rauscher, F.J. III. *The Wistar Institute, Philadelphia, PA 19104 *The Babraham Institute, Cambridge, UK.*

Transcriptional repressors of the C2H2 zinc-finger family contain a conserved repression domain called the KRAB domain. This domain has been shown to specifically bind the corepressor, KAP-1, which itself is capable of repressing transcription itself when tethered to a heterologous DNA-binding domain. KAP-1 and its mouse homolog, TIF-1$\beta$, are capable of binding HP1 family proteins. Members of the HP1 family share conserved domains termed chromo and chromoshadow domains and are associated with heterochromatic or silenced regions of chromatin. Thus one mechanism for KRAB-mediated repression of transcription may result from recruitment of HP1 proteins and formation of heterochromatin. We have begun a detailed analysis of the interaction between KAP-1 and human HP1$\alpha$ using purified proteins. Our analysis indicates that the central region of KAP-1 and the chromoshadow domain of HP1$\alpha$ are necessary and sufficient for direct binding between these proteins. We have begun experiments to determine which amino acid residues in the chromoshadow domain are important for binding to KAP-1 and whether these are conserved in the Drosophila HP1 protein. In addition, we have examined the activity of HP1 proteins expressed as fusions to a heterologous DNA-binding and observed that HP1 proteins are capable of repressing transcription 2 to 3-fold. This system will allow functional testing of the HP1 mutations and help to determine the extent to which transcriptional repression by KAP-1 may be mediated by HP1 factors.